

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CURSO DE GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

Aplicação da Teoria das Filas para melhoria do atendimento em
um consultório médico da cidade do Recife

Helder Lins de Carvalho
Orientador: Prof. Luciano Lins

RECIFE, Dezembro/2007



UNIVERSIDADE FEDERAL DE PERNAMBUCO

**DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO
CURSO DE GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO**

**APLICAÇÃO DA TEORIA DAS FILAS PARA MELHORIA
DO ATENDIMENTO EM UM CONSULTÓRIO MÉDICO
DA CIDADE DO RECIFE**

TRABALHO DE CONCLUSÃO DE CURSO DE GRADUAÇÃO
POR

HELDER LINS DE CARVALHO

Orientador: Prof^o Luciano Lins

RECIFE, DEZEMBRO / 2007

C331a

Carvalho, Helder Lins de.

Aplicação da Teoria das Filas para melhoria do atendimento em um consultório médico da cidade do Recife / Helder Lins de Carvalho. - Recife: O Autor, 2007.
52 folhas.

TCC (Graduação) – Universidade Federal de Pernambuco. CTG. Curso de Engenharia da Produção, 2007.

Inclui bibliografia e Apêndice.

1. Engenharia da Produção. 2. Teoria das Filas. 3. Consultório Médico – Atendimento. I. Título.

UFPE

658.5

CDD (22. ed.)

BCTG/2008-007

RESUMO

As filas de espera são um problema do mundo moderno. Sua presença em particular no setor de saúde está se intensificando e seu estudo, portanto, torna-se de grande relevância. O presente trabalho tem por objetivo a utilização da Teoria das Filas para avaliar e propor melhorias para os problemas de esperas de um consultório médico da cidade do Recife. Foi feita uma revisão bibliográfica sobre os modelos clássicos da Teoria das Filas, com seus principais elementos, assim como uma exposição da situação atual de esperas em serviços públicos de saúde e em consultórios particulares. Ao longo do trabalho, dados sobre as características do atendimento do consultório em estudo foram coletados e os parâmetros de performance foram estimados em cima dessas características. Esses dados foram então analisados graficamente para uso posterior no modelo de filas proposto e apresentação de possíveis soluções para o problema das esperas. Foram feitas recomendações de melhorias, como o balanceamento das chegadas dos clientes através de marcação de consultas ou orientação na chegada dos pacientes. Os resultados são mostrados através de simulação, que compara a situação atual do sistema com a situação de melhorias propostas, tendo sido constatados, principalmente, a diminuição do tempo de espera e do tamanho da fila e um ganho de disponibilidade do médico.

Palavras-chaves: Teoria das Filas, Consultório, Atendimento.

LISTA DE FIGURAS

Figura 1: Representação de um sistema de filas.....	6
Figura 2: (a) Sistema de fila única; (b) Sistema de filas múltiplas.....	8
Figura 3: Número de pacientes x Sexo dos pacientes	28
Figura 4: Número de pacientes x Faixa etária (anos)	28
Figura 5: Número de pacientes x Dia da semana	29
Figura 6: Número de pacientes x Faixa de horário de chegada.....	30
Figura 7: Número de pacientes x Tamanho da fila.....	31
Figura 8: Número de pacientes x Tempo de espera.....	32
Figura 9: Tempo de atendimento (minutos) x Sexo do paciente.....	32
Figura 10: Tempo de espera (minutos) x Faixa etária (anos)	33
Figura 11: Tempo de atendimento (minutos) x Faixa etária (anos)	33
Figura 12: Tempo de espera (minutos) x dia da semana	34
Figura 13: Chegada dos pacientes na quinta-feira.....	35
Figura 14: chegada dos pacientes na terça-feira.....	35
Figura 15: Tempo de atendimento x Dia da semana	36
Figura 16: Tamanho da fila x Dia da semana.....	36
Figura 17: Tempo de espera x Horário de chegada	37
Figura 18: Tempo de atendimento x Horário de chegada	38
Figura 19: Tamanho da fila x Horário de chegada	38
Figura 20: Tempo de atendimento x Presença de acompanhante.....	39
Figura 21: Tempo de atendimento x Pacientes de retorno	40
Figura 22: Distribuição dos tempos entre as chegadas.....	42
Figura 23: Distribuição dos tempos de atendimento	42

LISTA DE TABELAS

Tabela 1: Número de pacientes por data de coleta	26
Tabela 2: Número de pacientes por dia da semana	26
Tabela 3: Taxa de chegada	41
Tabela 4: Tempo de atendimento e taxa de atendimento	41
Tabela 5: Resultados do modelo M/M/1	43
Tabela 6: Quadro comparativo das simulações	44

SUMÁRIO

1	INTRODUÇÃO.....	1
1.1	Problemática.....	1
1.2	Justificativa.....	2
1.3	Objetivos.....	2
1.4	Metodologia.....	3
1.5	Organização do Trabalho.....	3
2	TEORIA DAS FILAS	4
2.1	Histórico	4
2.2	Conceito.....	4
2.3	Principais elementos de um sistema de filas	5
2.4	Notação.....	10
2.5	Medidas de desempenho de um sistema de filas	11
2.6	Modelos básicos de sistemas de filas	11
2.7	Aplicações	18
3	FILAS EM SERVIÇOS DE SAÚDE	20
3.1	Espera em Sistemas Públicos de Saúde.....	20
3.2	Espera em Consultórios particulares	21
4	RESULTADOS	23
4.1	Descrição do consultório	23
4.2	Coleta de dados.....	23
4.3	Análise de dados.....	26
4.4	Modelo de filas do consultório	41
4.5	Simulação	44
5	CONCLUSÕES	46
	REFERÊNCIAS BIBLIOGRÁFICAS	48
	APÊNDICE A	50
	APÊNDICE B.....	51

1 INTRODUÇÃO

1.1 Problemática

O mundo moderno, com todos seus benefícios, suas praticidades e seus aperfeiçoamentos, traz novos hábitos para a rotina da sociedade em geral. Com o desenvolvimento acentuado do mercado de serviços nas últimas décadas, a competitividade dos setores torna-se cada vez mais acirrada. Em meio a essa realidade, a figura do cliente surge como ponto central dos esforços e a sua satisfação deve ser buscada, mantida e superada sempre que possível.

Uma característica presente em grande parte do mercado de serviços é a presença de filas no sistema. As filas, dependendo de suas características, podem causar impaciência e nervosismo nas pessoas que necessitem enfrentá-las.

Quando as filas atingem o setor da saúde, várias outras características devem ser levadas em consideração. Uma longa espera numa fila para um serviço de saúde pode significar a complicação de casos clínicos ou em casos extremos a morte de pacientes, casos que infelizmente ainda se repetem no atual sistema público de saúde. Hospitais públicos superlotados e sem capacidade de atendimento da demanda são um triste exemplo que pode ser estudado com o uso da teoria das filas.

Outro problema que vem ganhando maiores proporções é o caso das longas esperas em consultórios médicos particulares, causadas pelas excessivas marcações, com o intuito de garantir o aproveitamento produtivo e a prevenção contra as desistências dos pacientes.

No presente trabalho teve-se a intenção de apontar soluções para problemas de atrasos referentes a um consultório médico da cidade de Recife, causados por uma grande demanda de pacientes no estabelecimento, buscando-se a melhoria do nível de serviço proporcionado.

1.2 Justificativa

É notória a relevância de um estudo do tema, visto que as longas esperas por atendimentos em consultórios podem significar uma percepção ruim por parte dos pacientes, criando a possibilidade desses pacientes não retornarem ao estabelecimento e ainda servirem de canal de veiculação de propaganda negativa do consultório para outras pessoas.

O estudo permite a obtenção de diversas melhorias no sistema como uma considerável redução do tempo de espera e conseqüente ganho de satisfação do paciente.

Possibilita-se ainda o planejamento dos atendimentos e do dimensionamento das marcações das consultas, além da determinação dos períodos críticos e balanceamento das chegadas. Indica também necessidades de expansões para atendimento da demanda e permite-se o dimensionamento do espaço físico referente ao local onde os pacientes vão aguardar (sala de espera do consultório).

1.3 Objetivos

1.3.1 Objetivo Geral

Solucionar o problema das esperas dos pacientes pelo atendimento no consultório fazendo uso de abordagem que permita o redimensionamento do sistema.

1.3.2 Objetivos Específicos

- Fazer uma revisão bibliográfica explicando-se os modelos básicos de filas e suas características principais.
- Fazer uma breve análise da situação das filas nos sistemas de saúde.
- Realizar coleta de dados e posterior análise com gráficos estatísticos (*histograma e box-plot*).
- Construir um modelo de filas para o consultório.
- Avaliar os impactos dos índices de desempenho obtidos no modelo através de uma simulação.

1.4 Metodologia

Recomenda-se uma análise gráfica de dados, em conjunto com conceitos e modelos básicos de filas para análise dos problemas de esperas em consultórios.

Fez-se uso de modelos clássicos da Teoria das Filas para a avaliação das várias características da fila como a taxa de chegada dos clientes, o tempo de atendimento do médico, o tempo médio de espera em fila de um cliente, o tempo médio de permanência do cliente no consultório, o tamanho médio da fila entre outras, além da construção de um modelo de filas do consultório.

Foi feita uma coleta de dados sobre o sistema através de formulários e a utilização de gráficos estatísticos (histograma e box-plot) para uma perfeita visualização desses dados e facilidade nas suas interpretações.

1.5 Organização do Trabalho

O trabalho foi organizado em cinco capítulos que serão resumidamente descritos a seguir.

O capítulo 1 traz a introdução do tema, sendo abordados a problemática, justificativa, objetivos e metodologia. O capítulo 2 trata da revisão bibliográfica sobre teoria das filas, sendo abordados conceitos e modelos básicos referentes ao tema, além de um breve histórico. O capítulo 3 comenta a situação atual das filas nos sistemas de saúde pública e nos consultórios particulares. O capítulo 4 traz a metodologia do trabalho, os resultados observados com a coleta de dados, análise gráfica dos dados, aplicação e estruturação em modelos de filas e simulação com os dados obtidos. O capítulo 5 é referente à conclusão do trabalho e traz sugestões para futuros desenvolvimentos a respeito do tema.

2 TEORIA DAS FILAS

2.1 Histórico

Segundo Fogliatti e Mattos (2007), os primeiros estudos sobre filas de espera datam de 1909 e são atribuídos ao matemático A.K. Erlang, considerado por vários estudiosos como o “pai” da Teoria das Filas por seu pioneirismo. Em 1917 Erlang Lançou a obra “*Solutions of Some Problems in the Theory of Probabilities of Significance in Automatic Telephone Exchanges*” que traz o fenômeno das chamadas aguardando linha para serem encaminhadas.

Posteriormente, vários modelos surgiram, como em 1936 com Adams e 1951 com Tanner para o tratamento do tempo de espera médio para pedestres atravessarem uma rua sem sinal; em 1953 com Everett para o caso de escoamento de barcos em portos; 1954 com Cobham para reparo de maquinarias; 1954 com Bailey para o direcionamento de pacientes na emergência de um hospital; 1962 com Morse e 1965 com Prabhu para o problema da formação de estoques; Em 1996 com Bitran e Morabito para processos de manufatura e Mendonça e Morabito em 2001 para um sistema médico de emergência. Também aplicada nos problemas computacionais com Ramamoorthy em 1965 e Courtois em 1977; aplicações em comunicações de computadores com Daigle em 1991 entre outros exemplos.

2.2 Conceito

A Teoria das Filas é uma ferramenta matemática para tratar de processos aleatórios, como atrasos e congestionamentos, em sistemas produtivos. Trabalha, portanto, com aproximações de situações do mundo real que envolvam espera.

“A teoria das filas consiste na modelagem analítica de processos ou sistemas que resultam em espera e tem como objetivo determinar e avaliar quantidades, denominadas medidas de desempenho, que expressam a produtividade/operacionalidade desses processos. Entre essas medidas, podem-se citar: número de elementos na fila, tempo de espera pelo atendimento e tempo ocioso dos prestadores de serviço.” (FOGLIATTI; MATTOS, 2007, p. 1).

Com o crescimento da economia e, mais especificamente, o crescimento do mercado de serviços, as filas de espera tornam-se cada vez mais presentes no cotidiano dos consumidores, causando insatisfação, nervosismo e o comprometimento do nível de serviço do estabelecimento. Existem estudos a respeito dos efeitos da espera no lado psíquico e emocional do homem, tendo o estresse como a principal consequência da vida moderna (NIE, 2000). Ao mesmo tempo, um serviço oferecido de forma dinâmica e com esperas reduzidas pode ser um fator importante em relação à competitividade e ao atendimento das necessidades dos clientes, obtendo-se ótimos níveis de satisfação. Entretanto, a definição de clientes para a teoria das filas é mais abrangente do que apenas consumidores humanos. Ela se estende por processos intercalados, funcionando como clientes e servidores, por cadeias logísticas de produção e fornecimento de bens, por sistemas de fluxo de informações como uma rede de telefonia entre outros vários exemplos de aplicações (JACOB, 2006; PINHEIRO, 2006; COSTA, 2007).

As filas se formam pela discrepância entre uma alta demanda (chegada de clientes) por determinado serviço e a capacidade de atendimento que um processo pode oferecer. A metodologia de filas auxilia no planejamento e no dimensionamento de um sistema para atender os clientes com presteza e sem atrasos, além de permitir o alcance de uma previsão de tempos de espera e a obtenção de soluções para a sua melhoria sem incorrer em um superdimensionamento (COSTA, 2007). Trata-se de um problema com visões conflitantes. De um lado os clientes, com seu enorme poder de influência e interessados em redução do tamanho da fila e do tempo de espera. Do outro a gerência, com objetivos de ganhos de produtividade, máxima utilização do serviço e diminuição da ociosidade (FOGLIATTI; MATTOS, 2007).

2.3 Principais elementos de um sistema de filas

Um sistema de filas é formado pelos seguintes componentes (COSTA, 2007):

- Processo de chegada;
- Tempo de serviço;
- Número de canais de serviço;
- Capacidade do sistema;

- População de usuários;
- Disciplina das filas.

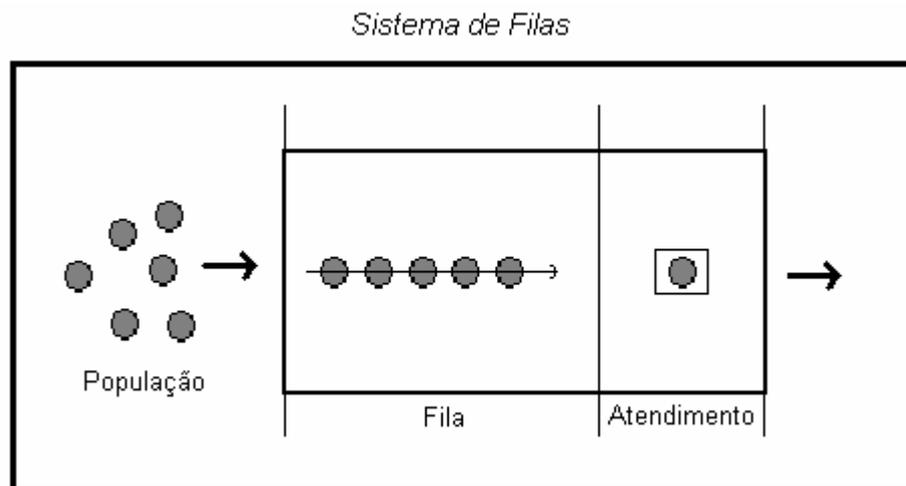


Figura 1: Representação de um sistema de filas

2.3.1 Processo de chegada

O processo de chegada estabelece um padrão de como os clientes chegam ao sistema. Nos casos comuns apresenta um comportamento estocástico, pois o tempo entre a chegada de clientes é uma variável aleatória. Torna-se fundamental o conhecimento da distribuição de probabilidade relacionada a esse tempo e a posterior determinação da taxa de chegadas (quantidade de clientes que chegam ao sistema por unidade de tempo).

A distribuição mais comum para a chegada dos clientes é a distribuição de Poisson na qual os tempos entre chegadas são exponencialmente distribuídos. (COSTA, 2007)

A distribuição exponencial tem algumas características importantes, como o tempo entre uma chegada e outra (intervalo entre chegadas) independente do tempo já decorrido. Assim, para um intervalo entre chegadas exponencialmente distribuído, teremos uma distribuição de Poisson para o número de chegadas no intervalo de tempo (MENDES, 2006).

Outros exemplos comuns de distribuições para tempo entre chegadas dos clientes são Erlang, hipereponencial e arbitrária.

É necessário saber se os clientes podem chegar ao sistema de forma simultânea, essa chegada é chamada de chegada batch ou em lotes. Nos casos em que ela for possível deve-se determinar distribuição de probabilidade do tamanho do batch.

A reação do cliente ao entrar no sistema pode alterar substancialmente a taxa de chegadas. Ele pode nem entrar no sistema por conta do tamanho da fila. Nesse caso é chamado de cliente decepcionado. Também pode desistir de permanecer na fila por conta da grande espera e sair definitivamente do sistema, ou trocar de fila quando o sistema possua filas paralelas. Nessas situações é então chamado de cliente impaciente.

A forma como o padrão das chegadas muda com o tempo também é importante. O padrão de chegadas é dito estacionário quando sua distribuição de probabilidade é independente do tempo, ou seja, não muda com o decorrer do tempo. E o padrão de chegadas não-estacionário é aquele que sua probabilidade não é independente do tempo (COSTA, 2007).

2.3.2 Tempo de serviço

Tempo de serviço é o tempo relacionado ao atendimento do cliente. É o intervalo de tempo que o cliente ficou realmente em atendimento. Aplicam-se as considerações sobre distribuição do processo de chegada ao tempo de serviço. Os serviços também podem ser simples ou em batch.

Servidor é a unidade de produção responsável pela realização, no todo ou em parte, do serviço.

O número de clientes na fila pode justificar uma mudança no tempo de serviço. Um servidor, por exemplo, pode acelerar o ritmo de trabalho quando perceber que a fila está grande ou aumentando, mas pode também se atrapalhar e ficar confuso, por conta do grande número de clientes esperando, comprometendo assim o tempo de serviço. Para essa situação dizemos que o sistema é dependente do estado, ou seja, que o sistema depende do número de clientes na fila. Por outro lado, o processo de atendimento em nada pode variar com o número de clientes na fila, nesse caso define-se o sistema como independente do estado.

Pode-se ainda relacionar o serviço com a sua variabilidade em relação ao tempo. Um servidor pode, por exemplo, obter ganhos de produtividade e melhorar sua eficiência de acordo com experiência adquirida ao longo do tempo, da mesma forma que pode diminuir seu rendimento em situações em que a atividade realizada acarrete cansaço ou fadiga com sua

continuidade. Nessas situações o sistema é dependente do tempo. Nos casos onde o sistema não sofre variações ao longo do tempo diz-se que ele é independente do tempo.

2.3.3 Número de canais de serviço

Esse parâmetro diz respeito ao número de servidores do sistema disponíveis para o atendimento dos clientes. Um processo que possui apenas um servidor é chamado de canal simples. Quando o mesmo possui mais de um servidor é chamado de multiservidor ou multicanal.

Quando existirem dois ou mais servidores a fila poderá ser organizada de duas situações. A primeira delas é um sistema de fila única no qual todos os servidores são demandados pela mesma fila (figura 2a). Como exemplo do primeiro caso pode-se citar a fila de pequenas compras de um supermercado, onde todos os clientes que se encaixarem no perfil aguardam na mesma fila pelo atendimento dos caixas-rápidos. O segundo caso de organização da fila é o de filas múltiplas, onde cada servidor possui sua própria fila, formando canais em paralelo em que cada canal opera de forma independente dos demais (figura 2b). Trazendo o exemplo do supermercado, os caixas para as compras de maior quantidade se encaixam no caso de filas múltiplas, com a formação de várias filas independentes das demais.

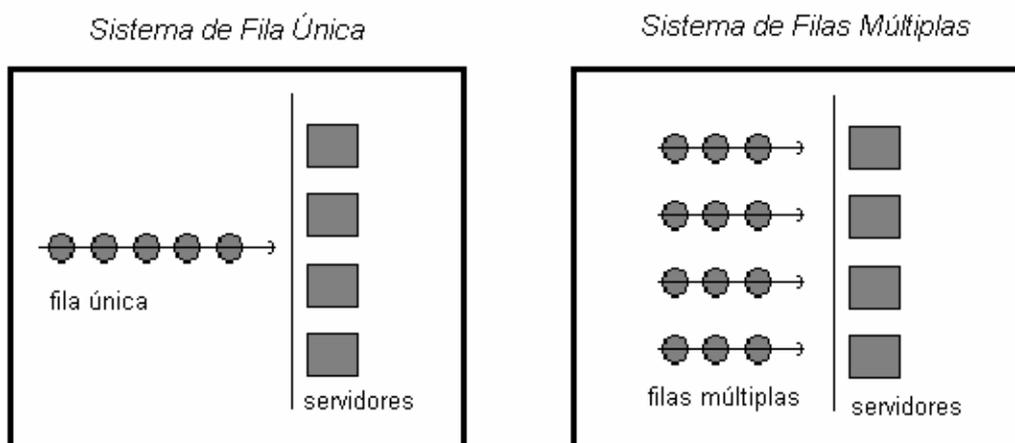


Figura 2: (a) Sistema de fila única; (b) Sistema de filas múltiplas

2.3.4 Capacidade do sistema

A capacidade do sistema diz respeito à quantidade de clientes que o sistema é capaz de suportar. É a quantidade de clientes esperando na fila mais a quantidade de clientes em atendimento.

A capacidade pode ser limitada. Dessa forma, quando o sistema estiver lotado, um novo cliente só poderá entrar depois de outro sair do sistema. A limitação pode ocorrer por razões de espaço físico ou custos para manter o cliente em fila por exemplo.

Pode existir também um sistema no qual todos os clientes que chegarem serão atendidos. Nesse caso diz-se que a capacidade do sistema é infinita.

2.3.5 População de usuários

Diz respeito ao número de clientes em potencial que poderão fazer parte do sistema de filas. A população de usuários pode ser finita, quando se conhece o número de potenciais clientes, ou infinita quando não se conhece a quantidade de clientes ou esse número é suficientemente grande para se utilizar o valor infinito do tamanho da população.

A população de usuários tem impacto direto na chegada dos clientes. Em uma população finita, a tendência é que a taxa de chegada diminua, à medida que os clientes vão entrando no sistema, pois a probabilidade de uma nova chegada diminui por conta do número de possíveis usuários diminuir.

2.3.6 Disciplina das filas

A disciplina das filas diz respeito a como os clientes serão selecionados para o atendimento depois da formação da fila. A disciplina mais tradicional é a FIFO (do inglês *First In, First Out* ou Primeiro a Entrar, Primeiro a Sair) onde é respeitada a ordem de chegada. O cliente que chegar primeiro será atendido com preferência sobre os demais. Como exemplo pode-se citar um atendimento bancário ou filas para entrada em um estádio de futebol. Mas existem outras disciplinas como a LIFO (do inglês *Last In, First Out* ou Último a Chegar, Primeiro a Sair), que é aplicada em alguns sistemas de armazenamento onde o último item guardado vai ser o primeiro a ser retirado.

Podem ocorrer também filas com prioridades, em que alguns clientes terão prioridade sobre os demais em duas situações: O primeiro caso é o preemptivo, em que o cliente com prioridade será atendido no momento de sua chegada, independentemente de outro cliente estar sendo atendido na hora. Nessa situação o atendimento atual será interrompido para o atendimento do cliente com prioridade, sendo retomado ao final do mesmo. Como exemplo pode-se citar uma unidade de atendimento de emergências de um hospital, onde alguns clientes precisam ser atendidos na frente de outros, por conta da gravidade de seus estados. O segundo caso é o não-preemptivo, no qual o cliente com prioridade vai para o início da fila, passando na frente dos demais, mas deverá aguardar o fim de um eventual atendimento já em andamento (COSTA, 2007).

2.4 Notação

Segundo Costa (2007), a notação mais comumente utilizada hoje em dia foi proposta por Kendall em 1953. Ela descreve um sistema de filas através da união das seis características dispostas acima e é descrita por símbolos da seguinte maneira: $A/S/m/K/N/Q$. Cada símbolo representa um parâmetro de filas. Assim temos “A” representando a distribuição dos tempos entre as chegadas, “S” representando a distribuição de probabilidade do tempo de serviço, “m” representando o número de canais de serviço, “K” representando a capacidade do sistema, “N” representando o tamanho da população de usuários e, por fim, “Q” representando a disciplina das filas.

Tem-se ainda:

- λ : taxa de chegadas dos clientes
- μ : taxa de atendimentos
- N : número de usuários no sistema
- N_q : número de usuários na fila
- T : tempo de permanência de um usuário no sistema
- T_q : tempo de permanência de um usuário na fila
- L : número médio de usuários no sistema
- L_q : número médio de usuários na fila
- W : tempo médio de permanência de um usuário no sistema

- W_q : tempo médio de permanência de um usuário na fila
- c : número de servidores
- ρ : taxa de utilização/ocupação do sistema
- $n(t)$: número de usuários no sistema no instante de tempo t

2.5 Medidas de desempenho de um sistema de filas

Quando se utiliza a teoria das filas para a análise e otimização de um processo e/ou para melhoria de seu nível de serviço devem-se buscar, segundo Fogliatti e Mattos (2007), os ganhos nas performances de alguns parâmetros, pode-se citar: número médio de usuários na fila (L_q) e número médio de usuários no sistema (L); tempo médio de permanência do usuário na fila (W_q) e tempo médio de permanência do usuário no sistema (W); probabilidade de se ter no máximo um número n_0 de usuários na fila, $P(N_q \leq n_0)$; probabilidade do tempo de espera na fila ser menor que um tempo t na fila, $P(T_q > t)$; probabilidade de se ter algum servidor ocioso, $P(N < c)$.

É possível também fazer o projeto do ambiente de espera do processo utilizando os valores estimados para o número médio de usuários na fila (L_q) (COSTA, 2007).

Tais parâmetros de desempenho dizem respeito aos objetivos dos clientes e/ou aos objetivos dos gerentes. O atendimento exclusivo a apenas uma das partes envolvidas pode comprometer como um todo devendo-se balancear os benefícios para ambos os lados.

2.6 Modelos básicos de sistemas de filas

Serão mostrados nessa seção os modelos basilares da teoria das filas que, com as devidas aproximações, permitem na maioria dos casos boas representações do objeto de estudo dentro do mundo real.

Antes da apresentação dos modelos torna-se necessária a explanação de dois conceitos como se seguem.

O primeiro deles é o de processos estocásticos. Como já falado anteriormente, os sistemas com tempos entre chegadas ou atendimento definindo variáveis aleatórias constituem

processos estocásticos. Segundo Fogliatti e Mattos (2007), estão associados a um processo estocástico dois tipos de espaço: o espaço de estados (E), que é o conjunto de valores que a variável aleatória $X(t)$ poderá assumir, e o espaço de parâmetros (U), que diz respeito ao conjunto de valores que a variável tempo (t) poderá assumir. Se o conjunto de valores que a variável aleatória pode assumir for um conjunto discreto, então o processo estocástico é chamado de cadeia estocástica. Por outro lado, se o conjunto de valores que a variável aleatória pode assumir for um conjunto contínuo o processo é chamado de contínuo.

O segundo conceito é o de processos markovianos. Um processo estocástico de parâmetro discreto $\{X(t), t > 0\}$ ou de parâmetro contínuo $\{X(t), t = \dots, -2, -1, 0, 1, 2, \dots\}$ é chamado de processo markoviano sempre que para uma informação do processo no presente, as informações passadas não interferem, de nenhuma maneira, na previsão desse mesmo processo no futuro, ou seja, o processo está com ausência de memória. Representando matematicamente teremos que para um conjunto qualquer dos n valores de tempo $t_1 < t_2 < \dots < t_n$ no conjunto faixa de tempo do processo, a distribuição de probabilidade condicional $X(t_n)$, dados os valores de $X(t_1)$, $X(t_2)$, $X(t_3)$, ..., $X(t_{n-1})$, dependerá somente do valor $X(t_{n-1})$.

O processo markoviano será uma cadeia de Markov sempre que o conjunto de valores da variável aleatória for um conjunto discreto, podendo variar entre cadeia de Markov e cadeia de Markov de parâmetro contínuo estando relacionadas respectivamente ao espaço de parâmetros com valores discretos e ao espaço de parâmetros com valores contínuos.

Quando o espaço de estado for um conjunto contínuo será possível o processo de Markov de parâmetro discreto e o processo de Markov de parâmetro contínuo.

Existe uma cadeia de Markov onde as únicas mudanças de estado são para seus vizinhos imediatos, ou seja, para um estado n , as únicas mudanças possíveis são para $n-1$ e $n+1$. Esse processo é denominado de processo de nascimento e morte. A mudança para o estado $n+1$ é chamada de nascimento e a mudança para o estado $n-1$ é chamada de morte (FOGLIATTI; MATTOS, 2007; COSTA, 2007).

2.6.1 Modelo M/M/1

Nesse modelo a distribuição dos tempos entre as chegadas e a distribuição de probabilidade do tempo de serviço são exponenciais. As chegadas e o atendimento

configuram um processo de nascimento e morte. A taxa de chegadas ao sistema (λ) e a taxa de atendimento (μ) são constantes. Existe apenas um canal de serviço. A capacidade do sistema é infinita (não existem restrições) e a disciplina das filas é a FIFO, ou seja, os clientes serão atendidos por ordem de chegada (FOGLIATTI; MATTOS, 2007).

No regime estacionário tem-se:

- Taxa de utilização do sistema

$$\rho = \frac{\lambda}{\mu} \quad (1)$$

- Probabilidade de se ter n clientes no sistema

$$P_n = \rho^n (1 - \rho) \quad (2)$$

- Probabilidade de se ter zero clientes no sistema

$$P_0 = 1 - \lambda / \mu \quad (3)$$

Para o cálculo das medidas de desempenho são necessárias as seguintes fórmulas:

- Número médio de usuários no sistema

$$L = \frac{\rho}{1 - \rho} \quad (4)$$

- Número médio de usuários na fila

$$L_q = \frac{\rho^2}{1 - \rho} \quad (5)$$

- Probabilidade de se ter mais do que k elementos no sistema

$$P(N \geq k) = \rho^k \quad (6)$$

- Tempo médio de espera em fila

$$W_q = \frac{\rho}{\mu - \lambda} \quad (7)$$

- Tempo médio de permanência no sistema: W

$$W = W_q + \frac{1}{\mu} = \frac{1}{\mu - \lambda} \quad (8)$$

- Probabilidade do tempo de espera ser maior do que um tempo $t > 0$

$$P(T_q > t) = \rho e^{-(\mu - \lambda)t} \quad (9)$$

Fórmulas de Little

Um resultado bastante relevante que não leva em consideração as distribuições de tempos entre chegadas e tempos de atendimento diz o seguinte: “O número médio de usuários num sistema é igual ao produto da taxa média de ingresso pelo tempo médio de permanência de um usuário no mesmo” (FOGLIATTI; MATTOS, 2007, p. 14). Esse resultado diz respeito à conhecida fórmula de Little que é matematicamente representada por:

$$L = E[\wedge]W, \quad (10)$$

onde $E[\wedge]$ é a taxa média de ingresso no sistema (FOGLIATTI; MATTOS, 2007).

Para o modelo M/M/1, da fórmula de Little e das suas derivações obtêm-se as seguintes fórmulas:

- Número médio de usuários no sistema

$$L = \frac{\lambda}{\mu - \lambda} \quad (11)$$

- Número médio de usuários na fila

$$L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} \quad (12)$$

- Tempo médio de permanência no sistema

$$W = \frac{1}{\mu - \lambda} \quad (13)$$

- Tempo médio de espera em fila

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)} \quad (14)$$

2.6.2 Modelo M/M/1/K

Nesse modelo a distribuição dos tempos entre as chegadas e a distribuição de probabilidade do tempo de serviço são exponenciais. As chegadas e o atendimento configuram um processo de nascimento e morte. Existe apenas um canal de atendimento e disciplina de fila é a FIFO. O diferencial desse modelo é que o sistema agora possui uma restrição na capacidade (K) repercutindo diretamente na taxa de chegadas dos clientes (λ). Dessa forma teremos dois valores para λ_n que irão depender do número de clientes no

sistema (n). Quando o valor de n estiver entre zero e o valor da capacidade do sistema (K) teremos a taxa de chegadas igual a λ , mas quando o valor de n se igualar ou ultrapassar o valor de K não serão permitidas chegadas no sistema, e o valor da taxa de chegadas será igual a zero (FOGLIATTI; MATTOS, 2007). Pode-se representar a taxa de chegadas por:

$$\lambda_n = \begin{cases} \lambda, & \text{se } 0 \leq n < K \\ 0, & \text{se } n \geq K \end{cases} \quad (15)$$

No regime estacionário tem-se:

- Taxa de utilização do sistema

$$\rho = \frac{\lambda}{\mu} \quad (16)$$

- Probabilidade de se ter zero clientes no sistema

$$P_0 = \begin{cases} \frac{1}{(K+1)}, & \text{se } \rho = 1 \\ \frac{(1-\rho)}{(1-\rho^{K+1})}, & \text{se } \rho \neq 1 \end{cases} \quad (17)$$

- Probabilidade de se ter n clientes no sistema

$$P_n = \begin{cases} \frac{1}{(K+1)}, & \text{se } \rho = 1 \\ \frac{(1-\rho)\rho^n}{(1-\rho^{K+1})}, & \text{se } \rho \neq 1 \end{cases} \quad (18)$$

Para o cálculo das medidas de desempenho são necessárias as seguintes fórmulas:

- Número de usuários no sistema

$$L = \begin{cases} \frac{K}{2}, & \text{se } \rho = 1 \\ \frac{\rho[1 + K\rho^{K+1} - (K+1)\rho^K]}{(1-\rho)(1-\rho^{K+1})}, & \text{se } \rho \neq 1 \end{cases} \quad (19)$$

- Número de usuários na filas

$$L_q = L - 1 + P_0 \quad (20)$$

- Tempo médio de permanência no sistema

$$W = \frac{L}{\lambda(1-P_K)} \quad (21)$$

- Tempo médio de espera na fila

$$W_q = W - \frac{1}{\mu} = \frac{L_q}{\lambda(1-P_K)} \quad (22)$$

- Probabilidade de se ter pelo menos k elementos no sistema ($k \leq K$)

$$P(N \geq k) = \begin{cases} \frac{(K+1-k)}{(K+1)}, & \text{se } \rho = 1 \\ \frac{(1-\rho^{K+1-k})\rho^k}{(1-\rho^{K+1})}, & k \text{ se } \rho \neq 1 \end{cases} \quad (23)$$

2.6.3 Modelo M/M/c

Como nos modelos anteriores, os tempos entre as chegadas e os tempos de atendimento seguem a distribuição exponencial, as chegadas e o atendimento configuram um processo de nascimento e morte e a disciplina das filas é a FIFO. A capacidade do sistema é infinita, mas agora existem c servidores no sistema, o que faz com que a taxa total de atendimento (μ_n) assuma dois valores distintos (FOGLIATTI; MATTOS, 2007). Pode-se representá-la matematicamente por:

$$\mu_n = \begin{cases} n\mu, & \text{se } 1 \leq n < c \\ c\mu, & \text{se } n \geq c \end{cases} \quad (24)$$

No regime estacionário tem-se:

- Taxa de utilização do sistema

$$r = \frac{\lambda}{\mu}; \quad \rho = \frac{r}{c} = \frac{\lambda}{c\mu} \quad (25)$$

- Probabilidade de se ter n clientes no sistema

$$P_n = \begin{cases} P_0 \frac{r^n}{n!}, & \text{se } 1 \leq n < c \\ P_0 \frac{r^n}{(c^{n-c})c!}, & \text{se } n \geq c \end{cases} \quad (26)$$

- Probabilidade de se ter zero clientes no sistema

$$P_0 = \left[\sum_{n=0}^{c-1} \frac{(r)^n}{n!} + \frac{(r)^c}{c!} \left(\frac{1}{1-r/c} \right) \right]^{-1} \quad (27)$$

Para o cálculo das medidas de desempenho são necessárias as seguintes fórmulas:

- Número médio de usuários na fila

$$L_q = P_0 \frac{cr^{c+1}}{c!(c-r)^2} \quad (28)$$

- Número médio de usuários no sistema

$$L = r + P_0 \frac{cr^{c+1}}{c!(c-r)^2} \quad (29)$$

- Tempo médio de espera em fila

$$W_q = P_0 \frac{r^c \mu}{(c-1)!(c\mu - \lambda)^2} \quad (30)$$

- Tempo médio de permanência no sistema

$$W = \frac{1}{\mu} + P_0 \frac{r^c \mu}{(c-1)!(c\mu - \lambda)^2} \quad (31)$$

2.6.4 Modelo M/M/c/K

Nesse modelo também ocorre de os tempos entre chegadas e os tempos de atendimento serem exponencialmente distribuídos, das chegadas e dos atendimentos configurarem um processo de nascimento e morte e da disciplina da fila ser a FIFO. Assim como no modelo anterior a taxa de atendimento pode assumir dois valores distintos, dependendo de quantos canais de atendimento são demandados. A capacidade limitada (K) também faz com que dois valores sejam atribuídos à taxa de chegada, dependendo se o sistema está lotado (atingiu a capacidade máxima de usuários) ou não (FOGLIATTI; MATTOS, 2007). Pode-se representar a taxa de atendimento e a taxa de chegada por:

$$\lambda'_n = \begin{cases} \lambda, & \text{se } 0 \leq n < K \\ 0, & \text{se } n \geq K \end{cases} \quad (32)$$

$$\mu_n = \begin{cases} n\mu, & \text{se } 1 \leq n < c \\ c\mu, & \text{se } c \leq n \leq K \end{cases} \quad (33)$$

No regime estacionário tem-se:

- Taxa de utilização do sistema

$$r = \frac{\lambda}{\mu}; \rho = \frac{r}{c} = \frac{\lambda}{c\mu} \quad (34)$$

- Probabilidade de se ter n clientes no sistema

$$P_n = \begin{cases} P_0 \frac{r^n}{n!}, & \text{se } 1 \leq n < c \\ P_0 \frac{r^n}{(c^{n-c} c!)}, & \text{se } c \leq n \leq K \end{cases} \quad (35)$$

- Probabilidade de se ter zero clientes no sistema

$$P_0 = \left[\sum_{n=0}^c \frac{(r)^n}{n!} + \frac{(r)^c}{c!} \sum_{n=1}^{K-c} \left(\frac{r}{c}\right)^n \right]^{-1} \quad (36)$$

Para o cálculo das medidas de desempenho são necessárias as seguintes fórmulas:

- Número médio de usuários na fila

$$L_q = \begin{cases} P_0 \frac{r^c (K-c+1)(K-c)}{c! \cdot 2}, & \text{se } r/c = 1 \\ P_0 \frac{r^{c+1} [(r/c-1)(K-c+1)(r/c)^{K-c} + 1 - (r/c)^{K-c+1}]}{c!c [1 - (r/c)^2]}, & \text{se } r/c \neq 1 \end{cases} \quad (37)$$

- Número médio de usuários no sistema

$$L = L_q + c + \sum_{n=0}^{c-1} (n-c)P_n \quad (38)$$

- Tempo médio de espera na fila

$$W_q = \frac{L_q}{\lambda'}, \text{ onde } \lambda' = \lambda(1 - P_k) \quad (39)$$

- Tempo médio de permanência no sistema

$$W = \frac{L}{\lambda'}, \text{ onde } \lambda' = \lambda(1 - P_k) \quad (40)$$

2.7 Aplicações

Nesse tópico serão mostradas algumas aplicações tendo caráter meramente exemplificativo da amplitude de aplicabilidade da ferramenta dentro do mercado e da sociedade como um todo, abordando inclusive os serviços governamentais.

As aplicações abrangem melhorias no oferecimento de alguns serviços como dimensionamento do número de caixas e terminais de auto-atendimento em bancos; formas de atendimento e disposição dos funcionários em uma lanchonete *fast-food*; dimensionamento e estabelecimento de prioridades em caixas de supermercado, como caixas exclusivos para gestantes e idosos ou caixas para pequenos volumes; redução do tempo de espera dos

pacientes e maior precisão da marcação de consultas em consultórios médicos particulares (objetivo de estudo do presente trabalho).

Podem ser objetos de estudo também alguns serviços de caráter público como o planejamento do processo de chegadas e envios de correspondências por parte dos correios e a avaliação do tempo de espera para conseguir um atendimento cirúrgico, como um transplante de órgãos, dentro do serviço público de saúde, avaliando também as conseqüências de seu mau dimensionamento (mortes, complicações nos quadros dos pacientes).

O planejamento de processos de produção, como, por exemplo, o de uma equipe de manutenção em uma fábrica, que trabalha com prováveis taxas de falhas no processo, onde o mesmo não pode parar por longos períodos, devido ao alto custo de espera.

São objetos de análise também qualquer forma de fluxo de tráfego como o trânsito de veículos nas grandes cidades, o congestionamento de aeroportos ou ainda em sistemas de comunicação como redes de telefonia e a própria internet.

Existem muitos outros exemplos de casos em que a teoria das filas pode ser utilizada com sucesso, o que vem a comprovar o poder que a ferramenta possui (DOY *et al.*, 2006; MORABITO & LIMA, 2000; NIE, 2000; SARMENTO JR *et al.*, 2005).

3 FILAS EM SERVIÇOS DE SAÚDE

A realidade de diversos estabelecimentos ligados a saúde, como consultórios, laboratórios e hospitais, está deixando grande parte dos clientes insatisfeita. O número de pacientes torna-se cada vez maior, mas não se percebe um atendimento médico capaz de prestar um serviço bem feito por completo, já que o tempo de espera para o atendimento, fator que vem ganhando importância na opinião dos clientes, tende a ser grande na maioria dos casos. Serão analisadas duas situações:

- O serviço público de saúde;
- Os consultórios particulares.

3.1 Espera em Sistemas Públicos de Saúde

O problema ganha maiores proporções quando a espera por um atendimento pode comprometer a integridade do paciente ou até mesmo por em risco a sua vida, como é o caso do atual serviço público de saúde, com estabelecimentos superlotados e incapazes de atender satisfatoriamente a população, especialmente a parcela mais carente.

Diversas variáveis influenciam a espera no sistema público. O paciente que procura um hospital público provavelmente já se encontra em situação de desconforto, dor, etc. Fatores como limpeza, organização, conforto acústico, boa iluminação, qualidade no atendimento tem repercussão direta na espera e o atual serviço público de saúde permanece deficiente em relação aos mesmos.

Nessas situações tenta-se estabelecer prioridades para os pacientes em estado mais grave, como acontece em atendimentos de emergência, mas a demanda é tão desproporcional à capacidade oferecida que pessoas acabam por ser gravemente prejudicadas por conta das esperas excessivas.

Por outro lado, aqueles que desejam um atendimento ou uma consulta de menor relevância encaminham-se para o estabelecimento desde a madrugada, antes do horário do funcionamento, na esperança de conseguir um atendimento após um período que chega a durar dias, semanas, meses ou até anos.

Verifica-se, portanto, a fragilidade extrema do setor e a urgente necessidade de investimentos e maior empenho governamentais na melhoria dessa situação (SARMENTO JR, *et al.*, 2005; CARVALHO, *et al.*, 2001).

3.2 Espera em Consultórios particulares

A espera de pacientes ligada ao atendimento em consultórios particulares assume diferentes pontos de vista. Atualmente, grande parte dos atendimentos em clínicas particulares é intermediada pelos planos de saúde. Como os planos estão remunerando os médicos com valores cada vez menores, estes são obrigados a aumentar o número de pacientes atendidos por dia para aumentar o aproveitamento do consultório.

Um fator de grande influência no planejamento das marcações é a desistência do paciente. Essa desistência ocorre quando um paciente marca uma consulta ou agenda sua presença, mas por motivos quaisquer deixa de comparecer ao consultório, o que pode causar a ociosidade do médico. Dessa maneira, para manter os consultórios sempre em funcionamento e aumentar sua produtividade, são marcadas mais consultas do que se é normalmente seria possível comportar. Assim consegue-se manter o estabelecimento sempre em atividade e com grandes índices de aproveitamento.

Mas essa marcação elevada acaba por gerar demandas acima da capacidade do estabelecimento, acarretando em esperas, atrasos e insatisfação, além de forçar o médico a atender os pacientes da forma mais rápida possível.

Do ponto de vista do paciente, vale ressaltar alguns pontos importantes em relação a sua percepção, são eles: o tempo de espera parece demorar mais a passar do que o tempo de atendimento; o tempo de espera parece andar mais rápido quando pequenos detalhes fazem o cliente pensar que o serviço está em andamento; esperas injustas são mais longas do que esperas justas, ou seja, quando um paciente é obrigado a esperar em situações que não o devia, a percepção da espera é maior do que a realmente experimentada; quanto maior o valor, que o paciente dá ao serviço, mais tempo ele está disposto a esperar. Existem alguns fatores do ambiente que também contribuem para a melhoria na percepção da espera como, por exemplo, a presença de revistas, jornais, música ambiente e televisão.

As esperas são desagradáveis e devem ser reduzidas ao máximo, pois o paciente deseja ser atendido num intervalo curto de tempo contado do momento do horário marcado para consulta até o momento da entrada no consultório.

Outro fator importante para o cliente é a forma como a consulta é direcionada. O tipo de atendimento oferecido pelo médico é bastante relevante, pois em diversas situações o paciente, além de passar por uma longa espera, é atendido às pressas e não tem correspondida sua expectativa que um bom atendimento compense a espera (MRTVI, 2007; CARVALHO, *et al.*, 2001).

4 RESULTADOS

4.1 Descrição do consultório

O consultório médico em estudo está situado na cidade de Recife/PE, no bairro de Areias, em uma avenida com grande fluxo de veículos. Funciona de segunda a sexta-feira no horário das sete horas às dezoito horas, mas cada médico possui seu horário específico. São atendidos pacientes conveniados e particulares, os quais são na sua maioria do próprio bairro ou dos bairros limítrofes.

São oferecidas quatro especialidades médicas, cada qual sendo representada por um profissional. São elas: clínica geral, otorrinolaringologia, pediatria e odontologia.

O clínico geral atende de segunda a sexta-feira das sete horas às oito e trinta e das dezesseis às dezoito e aos sábados das oito às doze. São atendidos em média dez pacientes por dia tendo como principais diagnósticos a hipertensão arterial, diabetes, viroses em geral, depressão, escabiose, candidíase, doenças gastrintestinais entre outros.

O otorrinolaringologista atende às terças e quintas-feiras das quatorze às dezesseis horas. São atendidos em média quinze pacientes por dia e os principais diagnósticos são sinusite, rinite, otite, resfriados, gripes, adenóide hipertrofiada e labirintite.

O pediatra atende apenas aos sábados das oito às doze horas. São atendidos em média dez pacientes tendo como diagnósticos principais gripes, resfriados, alergias, doenças respiratórias, diarreias, infecções, viroses e obesidade.

O cirurgião-dentista atende de segunda a sexta-feira das quatorze às dezessete horas e nas terças e quintas-feiras das oito às onze. São atendidos em média cinco pacientes em casa turno (manhã e tarde) e os principais atendimentos são relacionados a restaurações dentárias, exodontias, raspagem e polimento dentários, tratamento endodôntico, aplicação tópica de flúor, exames radiográficos e moldagens para prótese.

4.2 Coleta de dados

O trabalho se resumirá a fazer a análise de apenas uma especialidade devido aos prazos restritos para elaboração do sistema. A especialidade escolhida foi a otorrinolaringologia devido a melhor acessibilidade dos dados, a necessidade de o tempo de espera ser reduzido,

pois percebe-se a ocorrência de reclamações por parte dos pacientes relacionadas a demora, e a dificuldade que vem ocorrendo para a determinação da quantidade de pacientes a ser atendida por dia.

O atendimento do otorrinolaringologista foi dividido em dois tipos: primeira consulta, em que são recolhidos dados específicos e gerais do paciente como exames físicos, questionamentos sobre nome, idade, sexo, cor, endereço, telefone entre outros que são recolhidos pelo médico assim como a solicitação de exames complementares quando necessário e o diagnóstico junto com a prescrição de medicamentos e de possíveis tratamentos; e o retorno (segunda consulta), em que observam-se os resultados dos exames complementares e a evolução do tratamento diagnosticado. Geralmente a primeira consulta é mais detalhada, o que se espera resultar em uma duração superior a do retorno. As causas comuns de maiores esperas dos pacientes são devidas a elevada quantidade de atendimentos por dia e do grande número de chegada dos clientes comparado com o curto período que o médico permanece na clínica.

A recepção de pacientes funciona de duas maneiras: A maneira mais freqüente inicia-se com a ligação telefônica do paciente ou seu responsável para agendamento da consulta sem caráter de marcação de horário. A outra maneira é decorrente da chegada dos pacientes sem prévio aviso a recepção. Em ambas as formas os pacientes são atendidos de acordo com a ordem de chegada.

Os dados foram coletados através de um formulário que foi impresso para cada dia de coleta. O preenchimento foi feito pela secretária do turno, que foi a mesma para todos os dias da coleta. O formulário, em forma de planilha, possui os seguintes campos:

- Ordem de paciente: Diz respeito à ordem de chegada dos pacientes. Também serviu como uma forma de numerar e contabilizar os pacientes de cada dia.
- Sexo: Nesse campo é feita a simples diferenciação entre homens e mulheres, e para facilitar o preenchimento bastou apenas o preenchimento com “F” (feminino) ou “M” (masculino).
- Data de nascimento: O paciente quando chega ao consultório procura a secretaria e apresenta o seu documento de identificação. Através desse documento a secretária coletou as respectivas datas de nascimento.
- Convênio: Nesse campo foi feita a diferenciação dos pacientes de convênio e dos pacientes particulares sem fazer a distinção entre os convênios. Para facilitar o preenchimento foi colocado apenas “S” (sim, convênio) e “N” (não, particular).

- Horário de chegada: Foi colocado nesse campo o horário com a precisão de minutos que o paciente chegava ao consultório e procurava a secretaria.
- Horário de início de atendimento: Nesse campo foram preenchidos os horários também com a precisão de minutos que os pacientes se dirigiam para o consultório do médico.
- Horário do fim do atendimento: Foram coletados os horários com a precisão de minutos nos quais os pacientes saíam do consultório.
- Retorno: Foram coletados os dados para diferenciação entre pacientes de primeira vez, ou pacientes que são retorno de uma consulta prévia. Para facilitar foi feito o preenchimento com “S” (sim, retorno) ou “N” (não, primeira vez).
- Acompanhante: Nesse campo foi feita a distinção dos pacientes que entravam no consultório com acompanhantes diferenciando dos pacientes que entravam sozinhos. Foi preenchido com “S” (sim, pacientes com acompanhantes) e “N” (não, pacientes sem acompanhantes).

Foram coletados dados referentes a 118 pacientes, do dia 09 de outubro de 2007 ao dia 01 de novembro, durante três terças-feiras (09/10/2007, 16/10/2007 e 30/10/2007) e quatro quintas-feiras (11/10/2007, 18/10/2007, 25/10/2007 e 01/11/2007), no horário de atendimento da clínica, que em média era das treze horas, iniciando antes do horário previamente marcado devido à grande demanda (tabelas 1 e 2), às dezessete horas, terminando algumas vezes após o horário devido às esperas.

Tabela 1: Número de pacientes por data de coleta

Data	Dia	Pacientes
09/10/2007	Terça-feira	16
11/10/2007	Quinta-feira	13
16/10/2007	Terça-feira	22
18/10/2007	Quinta-feira	23
25/10/2007	Quinta-feira	14
30/10/2007	Terça-feira	21
01/11/2007	Quinta-feira	9
Total		118

Tabela 2: Número de pacientes por dia da semana

Dia	Média de Pacientes
Terça-feira	19,67
Quinta-feira	14,75

Não houve grandes problemas na coleta, pois existiu boa contribuição por parte das pessoas envolvidas e também por conta do formulário ser de fácil entendimento e rápido preenchimento.

4.3 Análise de dados

De posse de todos os dados coletados, foram criadas novas variáveis para utilizar na análise de dados. São elas:

- Idade: Com os dados das datas de nascimento foram obtidas as idades de todos os pacientes.
- Idade categorizada: As idades foram separadas em categorias. A categoria1 corresponde às pessoas com idade de zero a vinte e cinco anos, a categoria2 corresponde às pessoas com idade de vinte e cinco a cinquenta anos e, por fim, a categoria3, que corresponde às pessoas com idade superior a cinquenta anos.

- Dia: Foram separados os dados coletados na terça-feira e os dados coletados na quinta-feira.
- Chegada categorizada: Com os dados referentes aos horários de chegada ao consultório foram criadas categorias de acordo com a faixa de horário que os pacientes chegavam. A categoria1 corresponde aos pacientes que chegaram até às 14h, a categoria2 corresponde aos pacientes que chegaram até às 15h, a categoria3 corresponde aos pacientes que chegaram até às 16h e a categoria4 corresponde aos pacientes que chegaram até às 17h.
- Espera: Foram determinados os tempos de espera de todos pacientes utilizando o horário de início do atendimento menos o horário de chegada.
- Atendimento: Foram determinados os tempos de atendimento de todos pacientes utilizando o horário do fim do atendimento menos o horário de início do mesmo.
- Tempo entre chegadas: Foram calculados os intervalos de tempo entre duas chegadas consecutivas dos pacientes referentes a todas as chegadas.

Com todos os dados coletados e formatados, foi utilizado o software *Statistica 7.0* para a análise dos mesmos e busca por relações que poderiam afetar os parâmetros básicos de taxa de chegada e tempo de atendimento. Foram utilizados duas ferramentas estatísticas para a análise, o *box-plot* e o histograma.

O Box-Plot é um gráfico no formato de caixa, cujos limites são o 1º quartil e o 3º quartil, que representam 25% e 75% dos dados respectivamente. Esta caixa é dividida por uma linha, a mediana, que significa 50% dos dados. Existem também dois eixos, ou "bigodes", ligados à caixa estendendo-se aos extremos, isto é ao menor e ao maior valor dos dados, excluindo os valores discrepantes (*outliers*). De um extremo ao outro, temos o espalhamento dos dados. O Box-Plot é especialmente útil para mostrar a dispersão de um grupo de dados e as diferenças que existem entre grupos (CAMPOS, 2001).

O histograma, uma das sete ferramentas da qualidade, é uma representação gráfica formada por retângulos justapostos em que a base de cada um deles corresponde ao intervalo de classe e a sua altura corresponde à respectiva frequência. É um indicador da distribuição dos dados. Usados para análise de um grande conjunto de dados e permite a obtenção de resultados visuais bastante satisfatórios.

Foram então obtidas as considerações sobre o sistema em estudo. Em primeiro lugar, o número de pacientes do sexo feminino foi bem superior ao masculino, como mostra a figura 3.

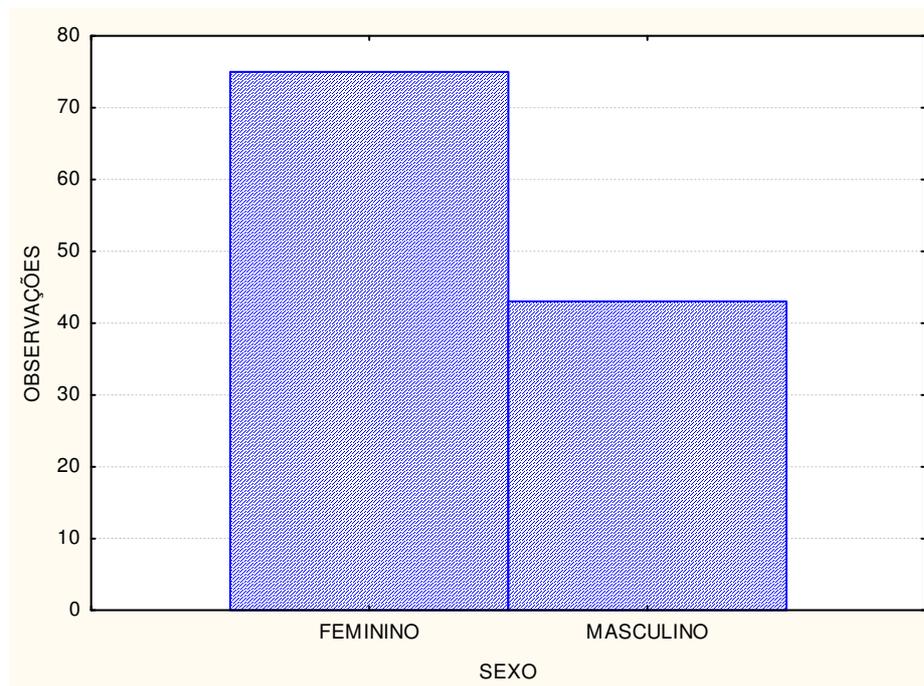


Figura 3: Número de pacientes x Sexo dos pacientes

A faixa etária dos pacientes foi bem variada. A maioria dos pacientes entre zero e dez anos, mas com grande número também relacionado à faixa etária de vinte anos aos cinquenta anos, como mostra a figura 4.

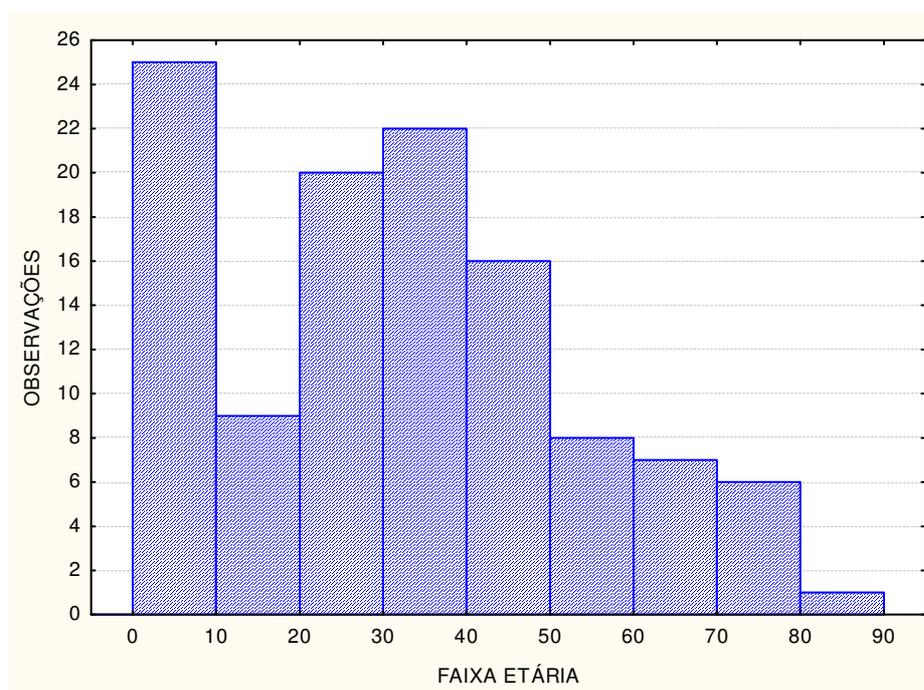


Figura 4: Número de pacientes x Faixa etária (anos)

O número de pacientes observados por dia da semana teve, por coincidência, valores iguais para a terça-feira e para a quinta-feira. Entretanto a coleta foi feita em três terças-feiras e quatro quintas-feiras, como mostra a figura 5, o que justifica o número maior de pacientes nas terças-feiras.

A chegada dos pacientes ao consultório se concentrou na faixa horária que vai das treze às quatorze. Percebeu-se que os pacientes chegavam ao consultório antes mesmo do médico, na tentativa de serem logo atendidos (figura 6).

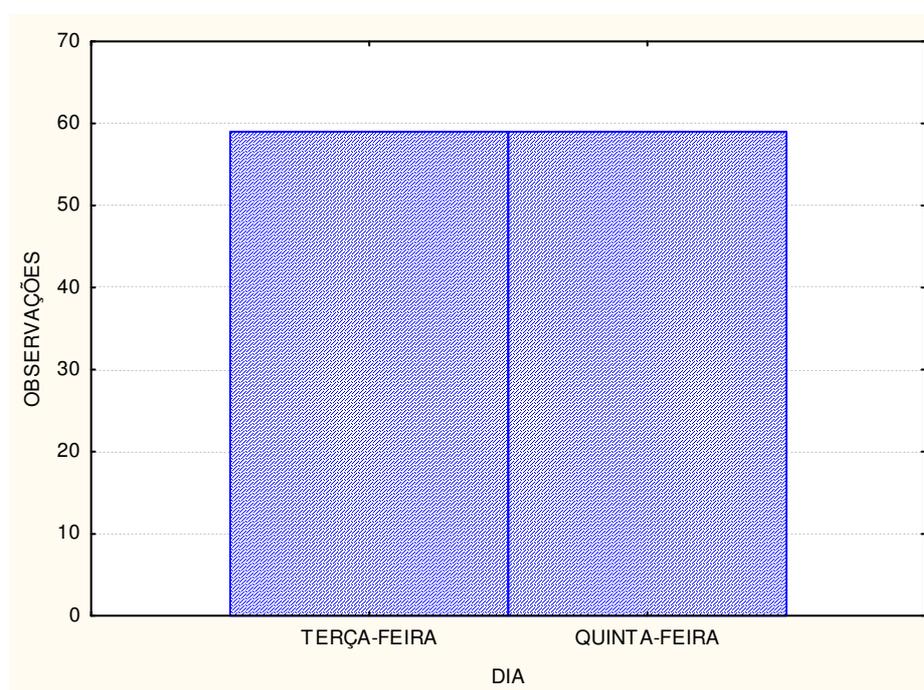


Figura 5: Número de pacientes x Dia da semana

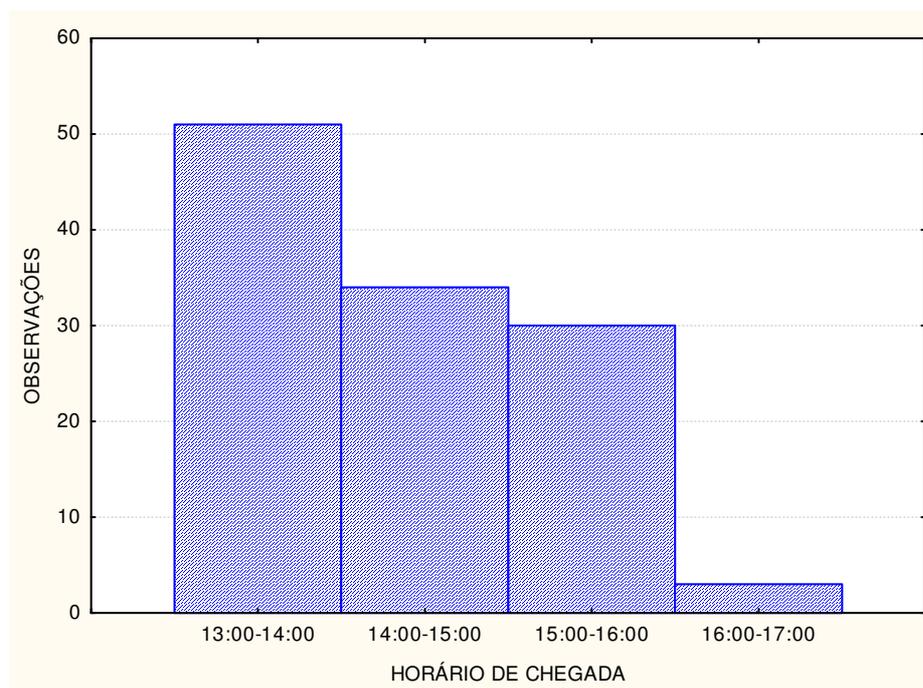


Figura 6: Número de pacientes x Faixa de horário de chegada

Avaliou-se o tamanho da fila que os pacientes encontravam quando chegavam ao consultório e confrontou-se o número de pacientes que chegavam com o número dos que se encontravam na fila. Percebeu-se que a maioria não encontrou muitos pacientes na fila, mas houve situações em que o tamanho da fila foi grande, com valores de oito, nove e até dez pacientes na fila (figura 7). Foram inclusive constatadas desistências por parte de alguns pacientes.

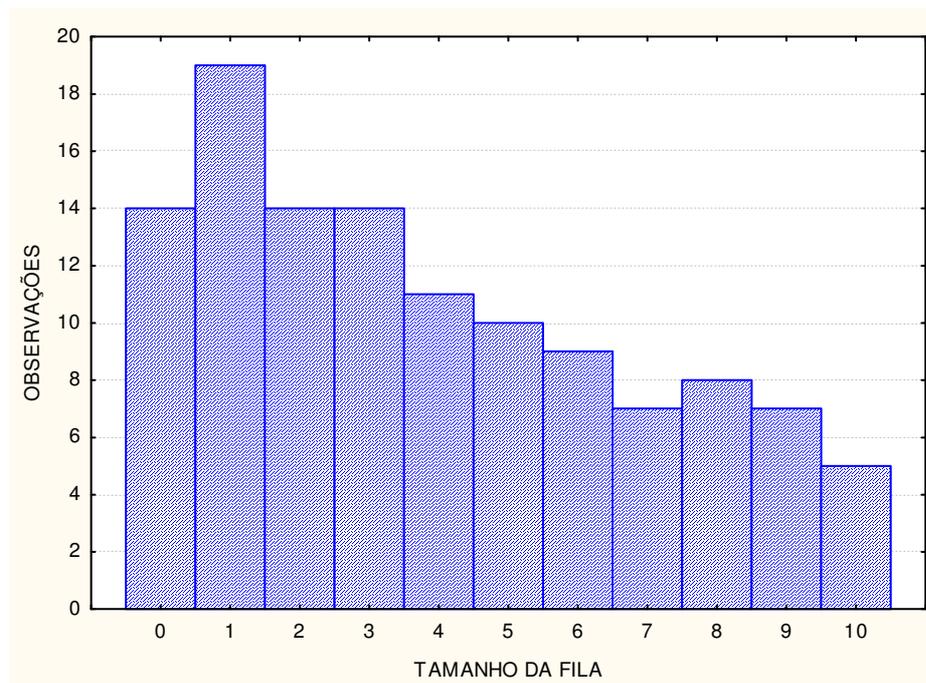


Figura 7: Número de pacientes x Tamanho da fila

Foi feita a análise também do número de pacientes em relação ao tempo que os mesmos esperavam no consultório e percebeu-se um número grande de pacientes esperando mais do que trinta minutos. Foram encontradas também ocorrências de pacientes esperando por mais de uma hora e, em casos remotos, esperas de aproximadamente duas horas (figura 8).

Verificou-se uma ligeira tendência de o tempo de atendimento dos pacientes do sexo masculino, com média de aproximadamente nove minutos e meio, ser menor em relação aos do sexo feminino, com média de dez minutos e meio (figura 9). Entretanto a relação entre tempo de atendimento e sexo não foi significativa.

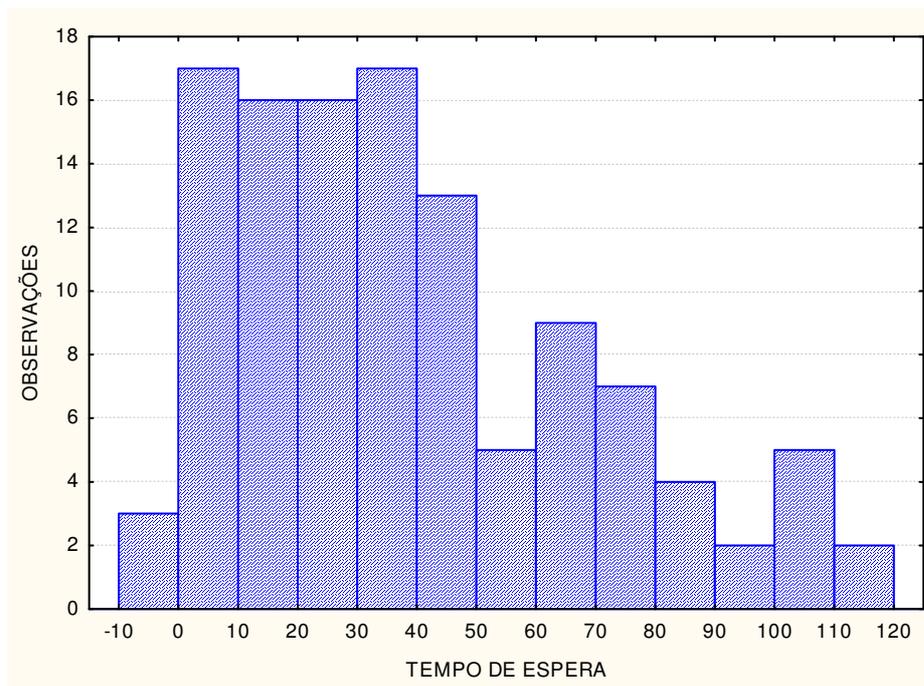


Figura 8: Número de pacientes x Tempo de espera

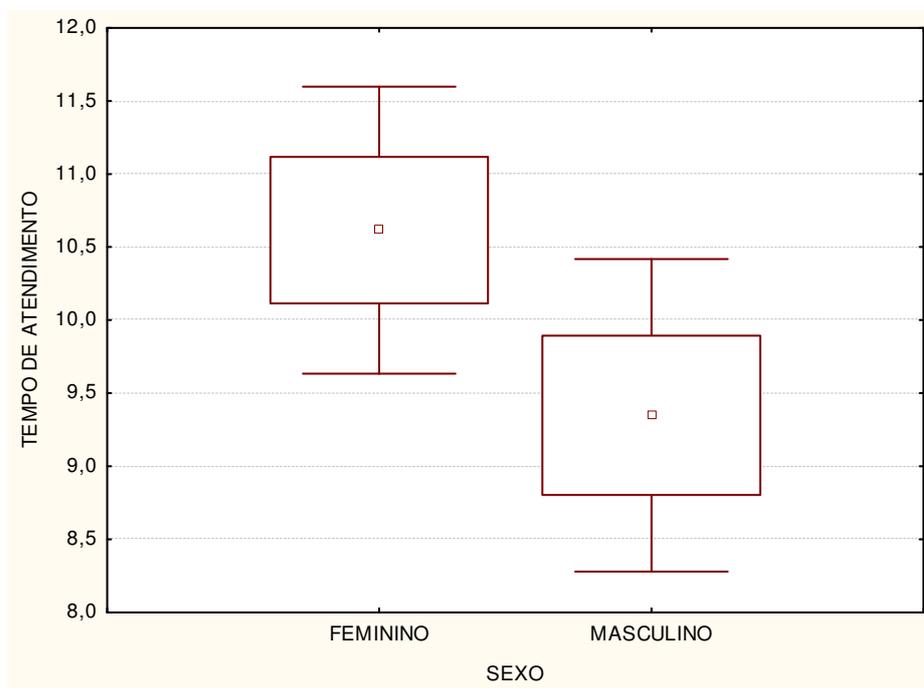


Figura 9: Tempo de atendimento (minutos) x Sexo do paciente

Verificou-se que o tempo total de espera e o tempo de atendimento não sofreram alterações relevantes quando analisados em relação à faixa etária dos pacientes (figuras 10 e 11).

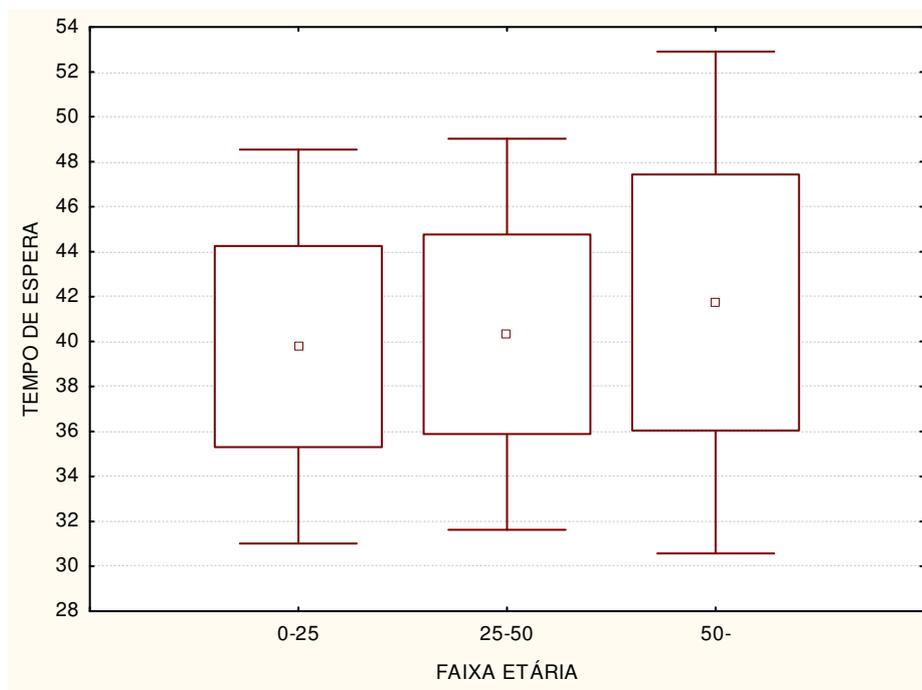


Figura 10: Tempo de espera (minutos) x Faixa etária (anos)

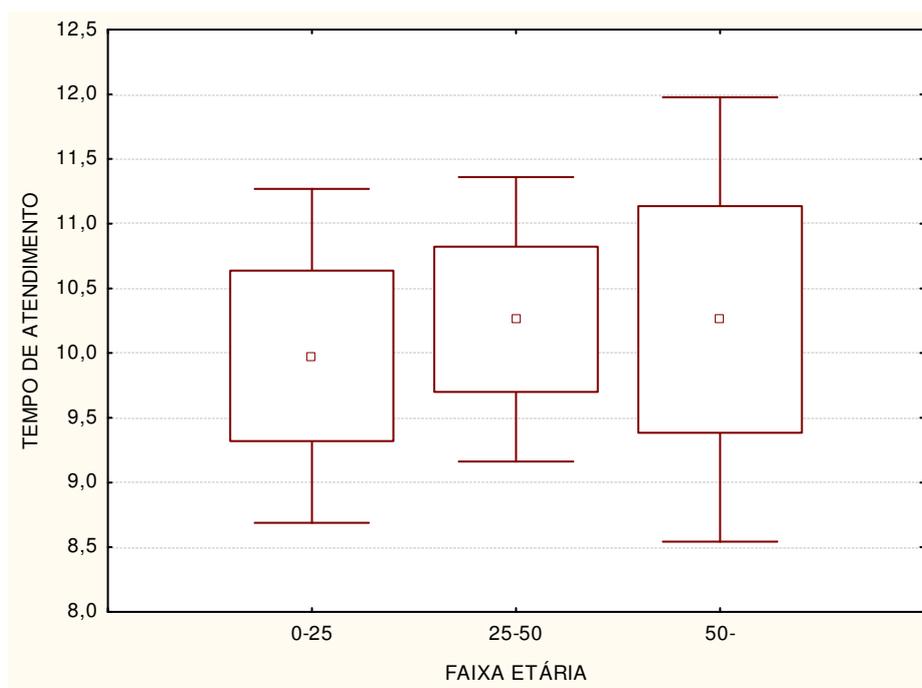


Figura 11: Tempo de atendimento (minutos) x Faixa etária (anos)

Quando foi feita a análise dos tempos de espera relacionados aos dias da semana percebeu-se que os pacientes da quinta-feira possuem uma tendência a esperar mais, média de quarenta e quatro minutos de tempo de espera, do que os pacientes da terça-feira, que esperam em média trinta e seis minutos (figura 12). Uma possível causa dessa espera maior na quinta-feira está relacionada à concentração da chegada dos pacientes no horário de início de funcionamento do consultório (das treze horas às quatorze horas e meia), enquanto na terça-feira a chegada dos pacientes está melhor distribuída (figuras 13 e 14).

O tempo de atendimento da quinta-feira não sofreu alterações relevantes quando comparado com o da terça-feira (figura 15). Constatou-se também uma tendência do tamanho da fila ser um pouco maior na quinta-feira, média de um pouco mais de quatro pessoas na fila, do que na terça-feira, media de pouco mais de três pessoas na fila (figura 16), provavelmente também devido à concentração das chegadas no início do funcionamento do consultório na quinta-feira.

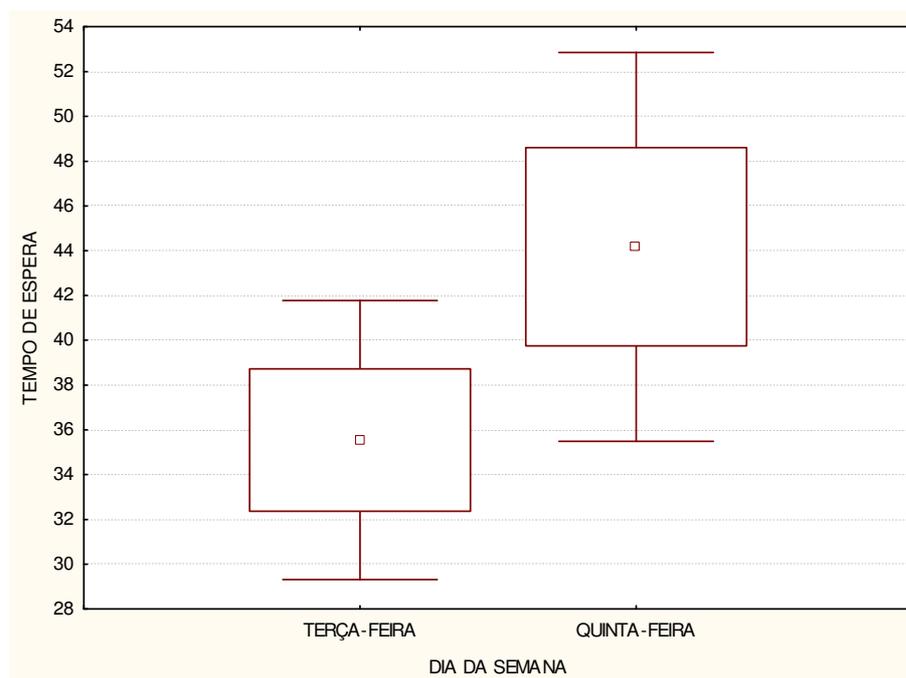


Figura 12: Tempo de espera (minutos) x dia da semana

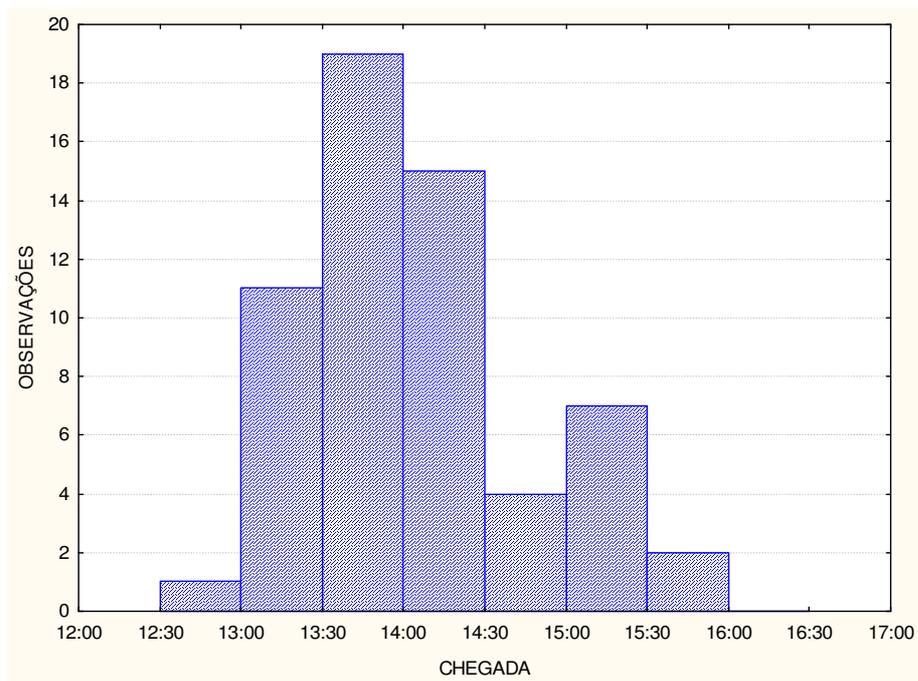


Figura 13: Chegada dos pacientes na quinta-feira

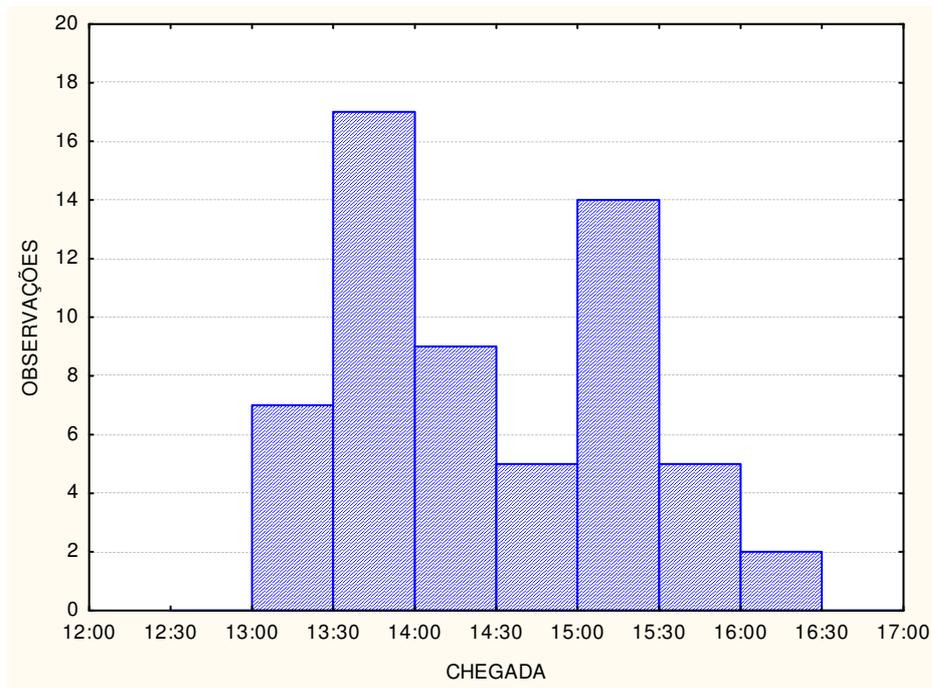


Figura 14: chegada dos pacientes na terça-feira

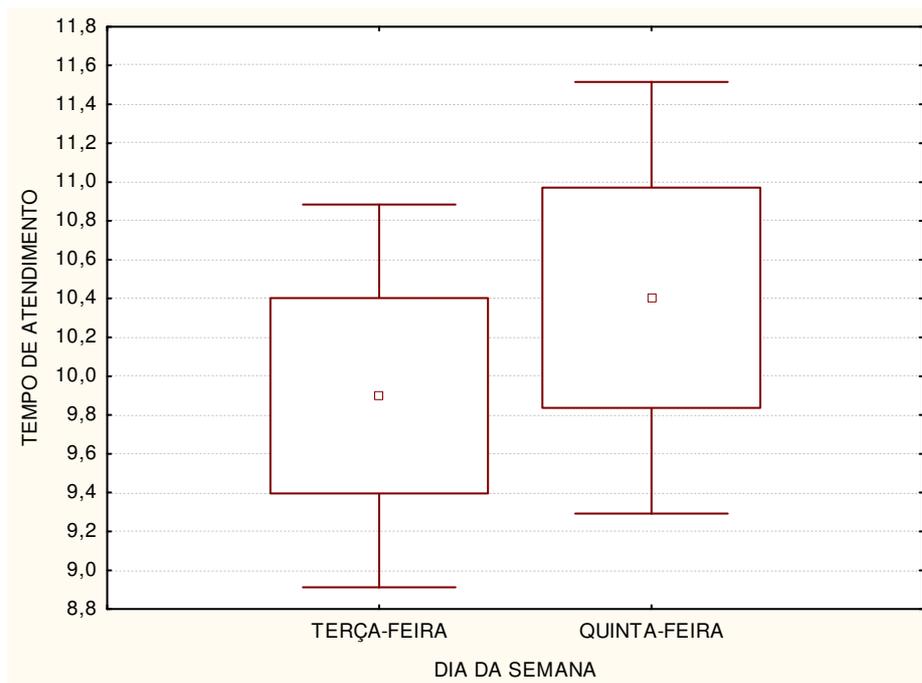


Figura 15: Tempo de atendimento x Dia da semana

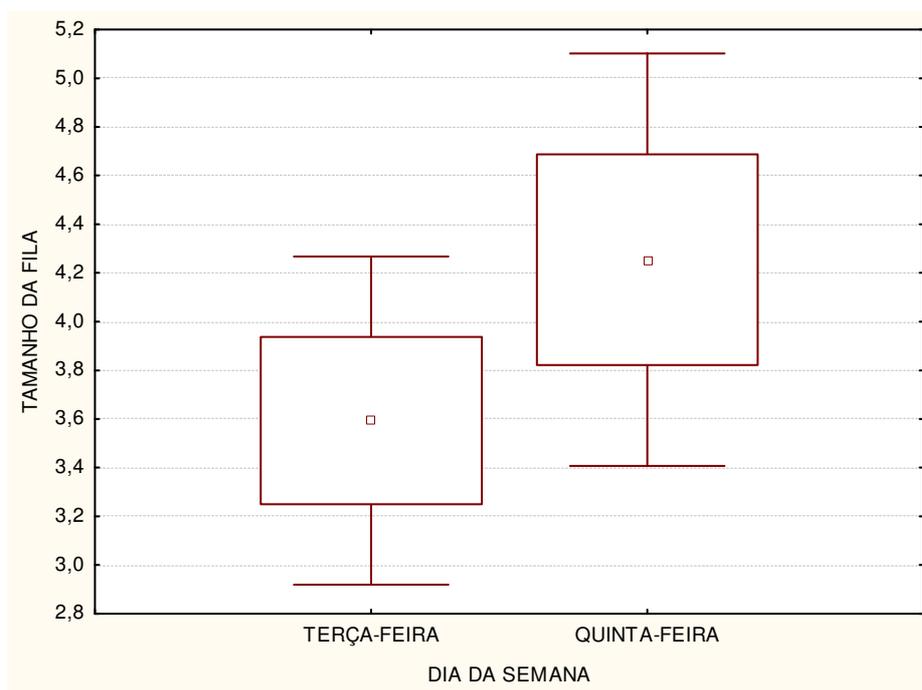


Figura 16: Tamanho da fila x Dia da semana

Analisando os dados quanto à faixa de horário de chegada, percebe-se uma tendência de que um paciente que chega ao consultório na faixa das quatorze horas às quinze horas tenha seus tempos de atendimento e espera e tamanho da fila maiores que os pacientes que chegam

nas outras faixas, principalmente quando comparados com os resultados dos pacientes que chegam na faixa das dezesseis às dezessete horas, que possuem tempos de espera e tamanho da fila menores em relação aos demais. Conclui-se então que o processo de chegada se concentrava no início do atendimento, a partir das treze horas e também na faixa que ia das quatorze às quinze horas. No final do expediente o consultório já estava mais vazio e os pacientes que chegavam na faixa das dezesseis às dezessete eram atendidos com mais rapidez por conta da pequena fila presente (figuras 17, 18, 19).

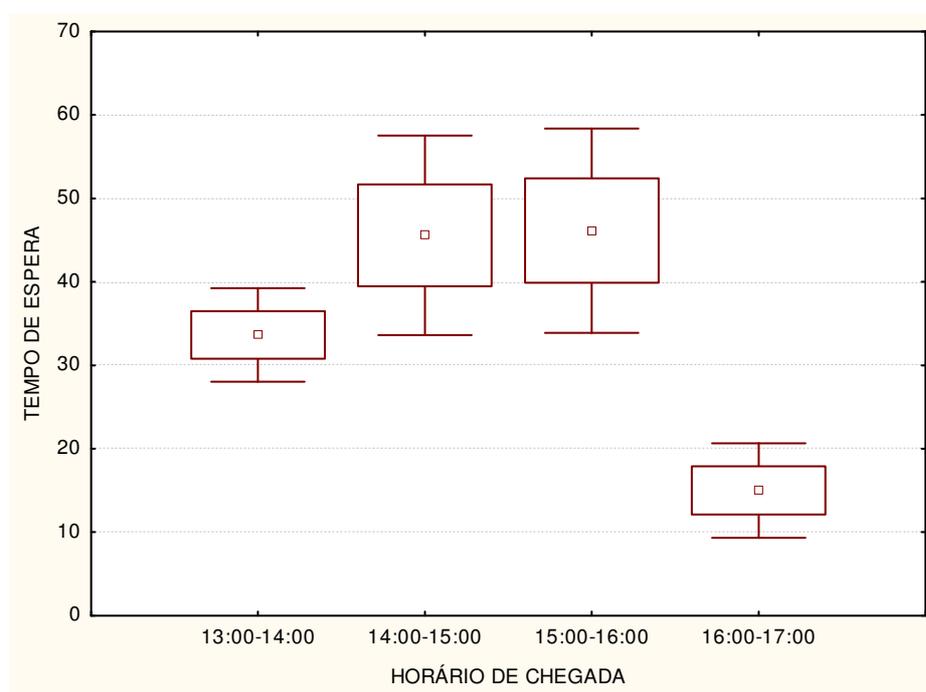


Figura 17: Tempo de espera x Horário de chegada

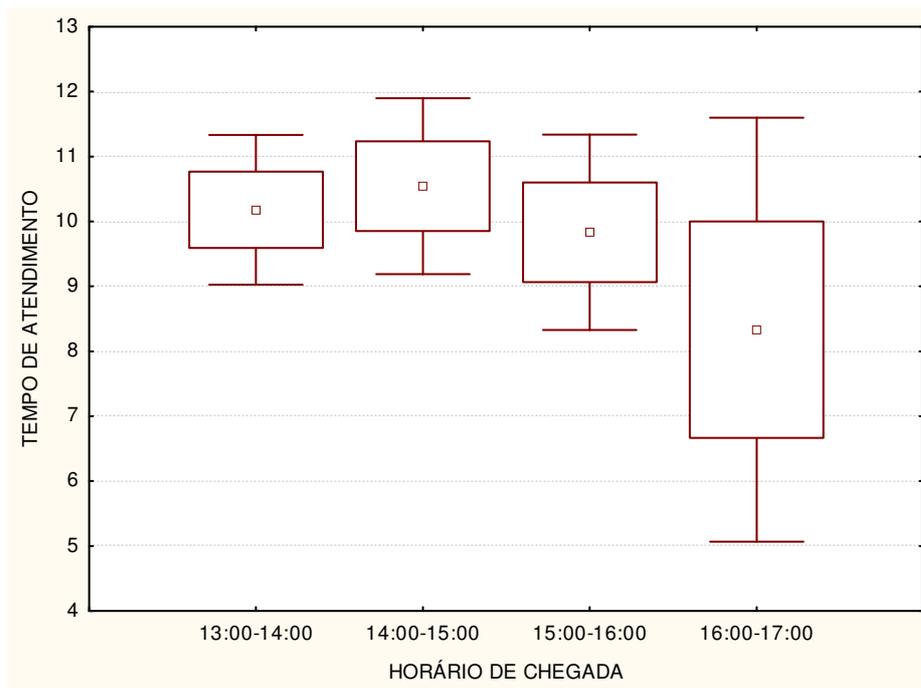


Figura 18: Tempo de atendimento x Horário de chegada

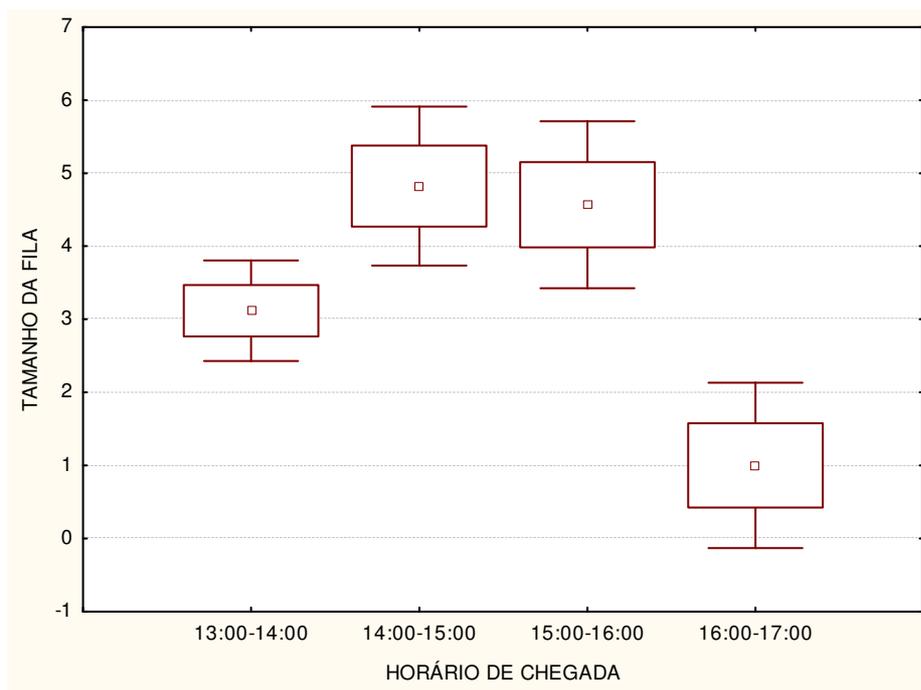


Figura 19: Tamanho da fila x Horário de chegada

Analisou-se uma tendência de que o tempo de atendimento dos pacientes que iam para a consulta médica na presença de um acompanhante era um pouco maior dos que eram atendidos sozinhos (figura 20). Essa tendência pode ser talvez explicada por conta de o

acompanhante contribuir com informações durante a consulta. Entretanto, mesmo que uma diferença seja constatada, ela deverá ser pouco expressiva.

A quantidade de dados coletados para a confrontação do tempo de atendimento dos pacientes em primeira consulta com os pacientes de retorno não foram suficientes para considerações mais expressivas. Contudo é possível perceber uma leve tendência de o tempo de atendimento dos pacientes de retorno ser um pouco menor que o dos pacientes de primeira vez, como mostra a figura 21, mas não deve ser muito discrepante.

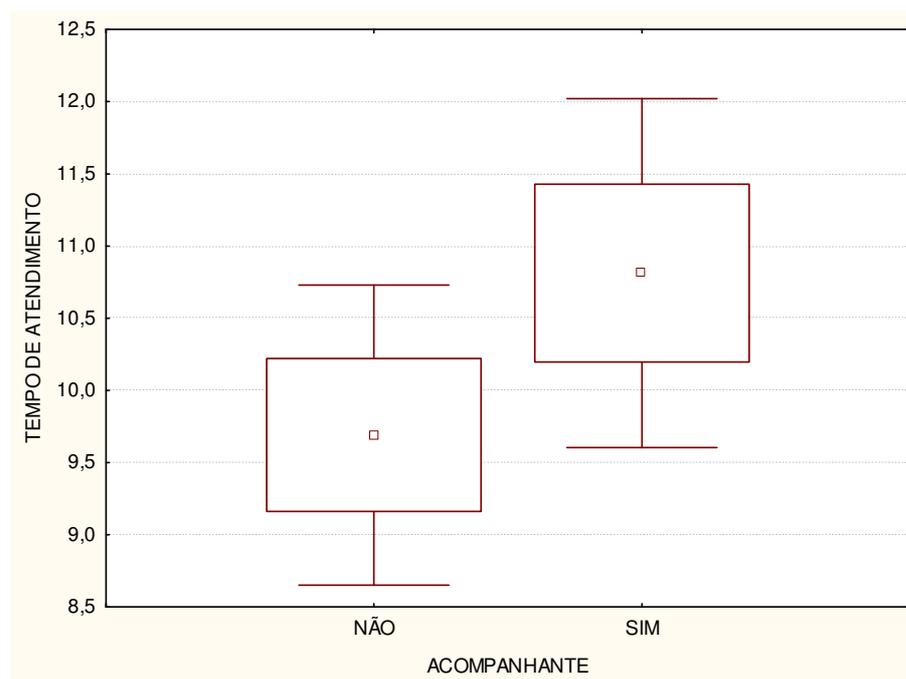


Figura 20: Tempo de atendimento x Presença de acompanhante

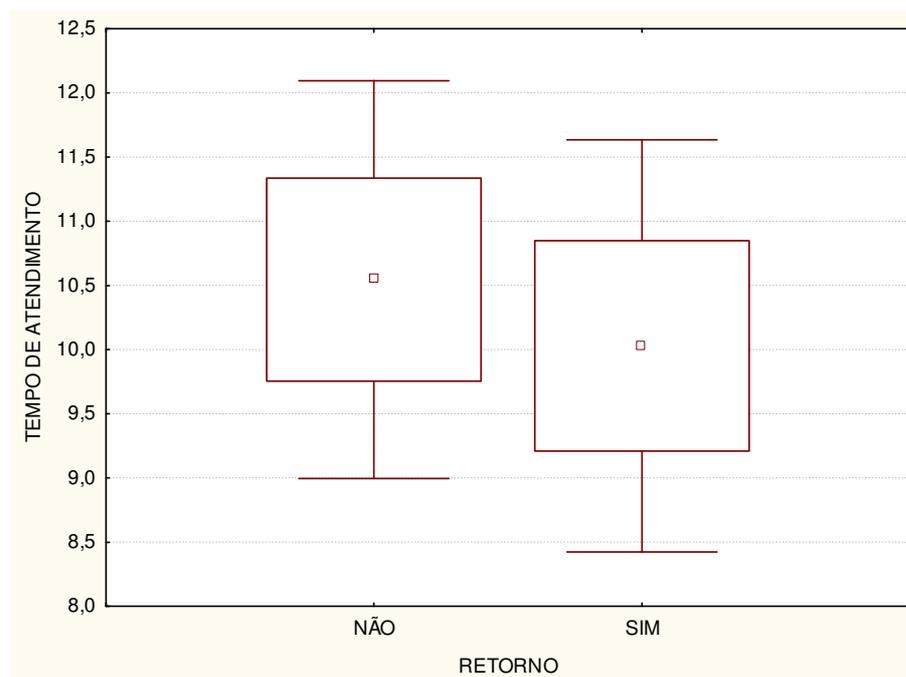


Figura 21: Tempo de atendimento x Pacientes de retorno

Percebe-se então que os problemas de esperas e de tamanho da fila são bastante relevantes (figuras 7 e 8) e que sua possível causa seja a tendência que os pacientes possuem de concentrar suas chegadas no horário de início de atendimento (figura 6). Foi também constatado, durante a coleta de dados, que existem desistências causadas pela longa espera decorrente de um grande número de pacientes.

Em cima dos dados analisados é proposto o modelo de filas do consultório.

4.4 Modelo de filas do consultório

Foram analisados os dados referentes às chegadas dos pacientes em cada faixa horária de todos os dias do atendimento e posteriormente a consolidação da taxa de chegada (pacientes por hora) da quinta-feira e da terça-feira. Ao final foi tirada a média dos dois dias, pois calcular o modelo para os dias separados não seria possível, visto que a fila da terça-feira não seria estável (taxa de chegada > taxa de atendimento). As informações foram consolidadas na tabela 3.

Tabela 3: Taxa de chegada

Taxa de chegada (pacientes/hora)	Terça-feira	Quinta-feira
13h-14h	8,00	6,75
14h-15h	4,00	5,50
15h-16h	6,67	2,50
Média	6,22	4,92
Média Geral	5,57	

Não foi identificada diferença significativa entre o tempo de atendimento na terça-feira e na quinta-feira (figura 15). Sendo assim, foi calculada uma taxa de atendimento única para os dois dias (tabela 4).

Tabela 4: Tempo de atendimento e taxa de atendimento

Tempo de atendimento (minutos/paciente)	Taxa de atendimento (pacientes/hora)
9,89	6,06

De posse dos parâmetros essenciais foi proposto o modelo M/M/1, com 1 servidor (o médico), sem restrições de capacidade e os clientes sendo atendidos pela ordem de chegada (FIFO). Nesse modelo a distribuição dos tempos entre as chegadas e a distribuição de probabilidade do tempo de atendimento devem ser exponenciais. No consultório, a distribuição dos tempos entre as chegadas teve um comportamento exponencial, como mostra a figura 22. A distribuição de probabilidade do tempo de atendimento não teve, exatamente, um comportamento exponencial (figura 23), mas foi admitida para poder fazer uso do modelo proposto, pois foi imposta uma limitação na simulação realizada.

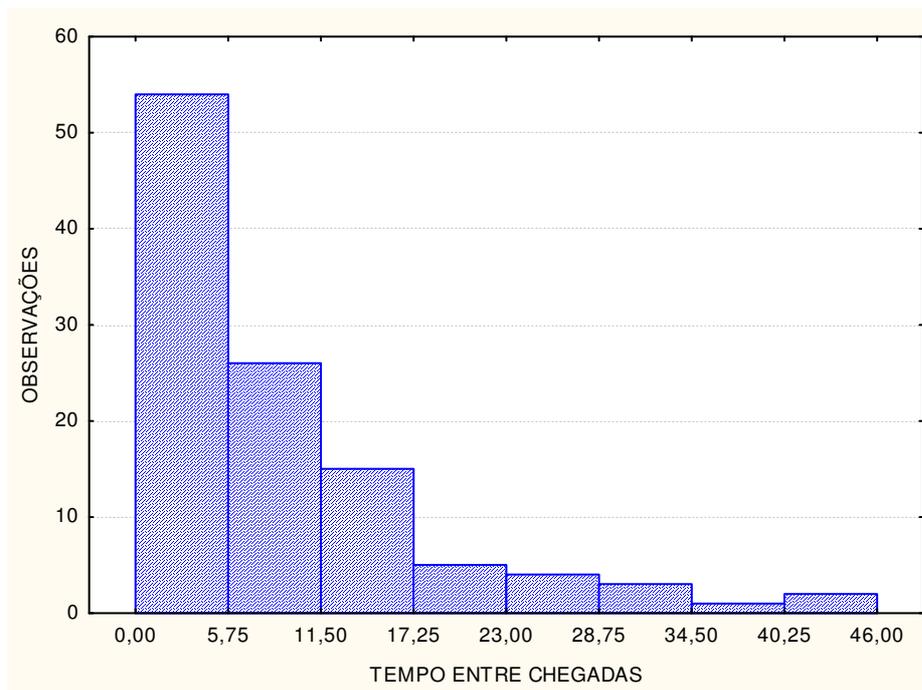


Figura 22: Distribuição dos tempos entre as chegadas

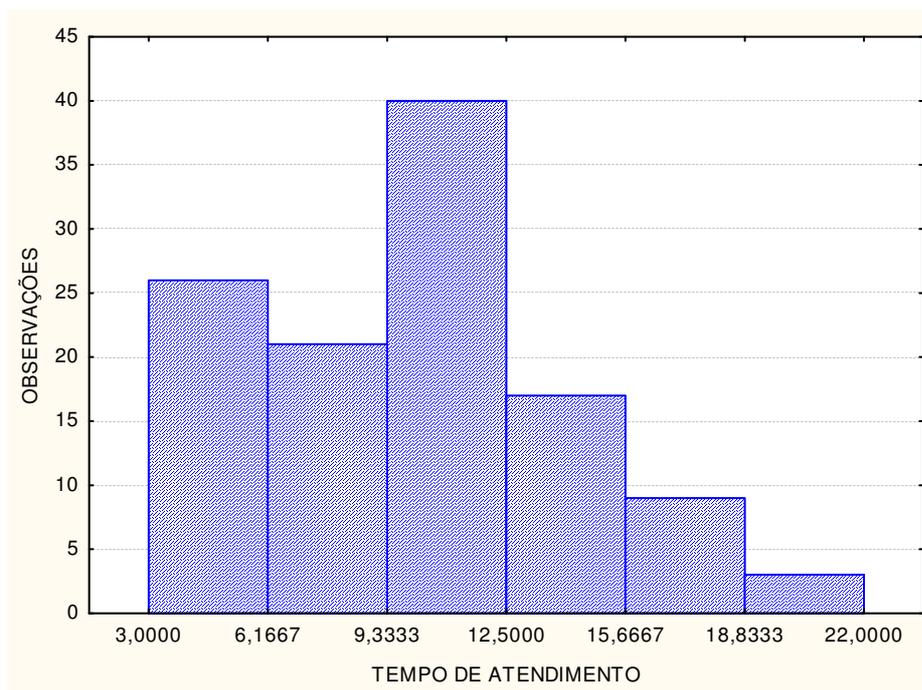


Figura 23: Distribuição dos tempos de atendimento

Aplicando o modelo M/M/1 no consultório com os valores de $\lambda = 5,57$ pacientes/hora e $\mu = 6,06$ pacientes/hora tem-se:

Tabela 5: Resultados do modelo M/M/1

Taxa de chegada	λ	5,57 pacientes/hora
Taxa de atendimento	μ	6,06 pacientes/hora
Taxa de utilização do consultório	ρ	0,92
Número médio de pacientes na fila	L_q	10,45 pacientes
Número médio de pacientes no sistema	L	11,37 pacientes
Tempo médio de permanência de um paciente na fila	W_q	1,88 horas
Tempo médio de permanência de um paciente no sistema	W	2,04 horas

Deve-se salientar que os valores encontrados representam o sistema no seu estado permanente e que devido à restrição de horário (a clínica, normalmente, só atende até as 16 horas) o sistema de filas do consultório fica apenas em estado transitório, ou seja, não atinge o estado permanente.

4.4.1 Sugestões de melhoria

Verificou-se que a causa dos problemas das esperas e tamanho da fila foi a concentração das chegadas dos pacientes no início de horário do consultório. Sendo assim, as propostas de melhoria são:

- Balancear a chegada dos clientes através de um sistema de marcação de consultas, distribuindo as chegadas ao longo do horário de funcionamento da clínica.
- Balancear as chegadas com orientações aos clientes sobre os melhores horários para as chegadas de forma a minimizar as esperas
- Iniciar os atendimentos 30 minutos após as chegadas dos primeiros clientes, para evitar a ociosidade do médico no início do atendimento.
- Impor uma restrição de horário para as chegadas às 15:30 horas, para que o médico, em média, não ultrapasse o horário de funcionamento do consultório, que é às 16:00 horas.

4.5 Simulação

Foram realizadas várias simulações em cima dos dados coletados e confrontadas duas situações: a realidade atual e a proposta de melhoria do consultório com o balanceamento do tempo de chegadas.

Na proposta atual foi usado o valor de taxa de chegada inicial igual a 9 pacientes por hora (representando a concentração da chegada no início do horário do consultório) e foi estipulado um decrescimento desta taxa em 6% para representar a diminuição da taxa de chegadas com o decorrer do horário do atendimento, já que percebeu-se que a chegada dos pacientes se concentrava no início do atendimento e diminuía com o decorrer do tempo. Portanto:

$$\lambda_{j+1} = (1 - d)\lambda_j, \quad j = 1, 2, \dots, m - 1 \quad (41)$$

onde:

- λ_j : taxa de chegada do j-ésimo paciente
- d : decrescimento da taxa de chegada
- m : número total de pacientes em um dia

Para a proposta de melhoria foi usada a taxa de chegada balanceada entre os dois dias, uniforme em relação ao tempo, igual a 5,57 pacientes por hora, ou seja, taxa de chegada de pacientes inicial de 5,57 pacientes por hora e decrescimento igual a zero. Para taxa de atendimento foi usado o mesmo valor para as duas situações, 6,06 pacientes por hora.

Com as simulações, cujos dados completos encontram-se no apêndice B, foram obtidos os resultados que são mostrados na tabela 6.

Tabela 6: Quadro comparativo das simulações

Elementos	Situação Atual	Situação proposta
Número de pacientes por dia	17,7	15,9
Tempo entre chegadas (minutos)	6,73	9,09
Tempo de atendimento (minutos)	7,64	7,71
Tempo médio de espera (minutos)	31,09	20,25
Tamanho médio da fila (pacientes)	4,01	2,40
Taxa de utilização do consultório	0,97	0,89
Horário do fim dos atendimentos	15:50	15:49

Percebe-se que o número esperado de pacientes por dia foi próximo nas duas situações (17,7 dias e 15,9 dias). O tempo entre chegadas sofre alterações, pois na situação proposta ele aumentaria em média para 9,09 minutos. O tempo esperado de atendimento não sofre alterações expressivas (7,64 minutos e 7,71 minutos).

As simulações mostram que um balanceamento da taxa de chegadas pode trazer melhorias significativas no tempo médio de espera (redução de 31 minutos para 20 minutos) e no tamanho médio da fila (redução de 4,01 pacientes para 2,4 pacientes). Essa diminuição do número esperado de pacientes na fila pode diminuir a incidência das desistências percebidas durante a coleta de dados. Percebe-se que a taxa de utilização diminui com o balanceamento das chegadas (de 97% para 89%), o que permite que o médico dedique um pouco mais de tempo para os atendimentos, melhorando o *rapport* (“entrar em sintonia”) entre médico e paciente (SILVA, 2001). Foi observado também que o balanceamento das chegadas pode gerar ociosidade no início do atendimento. Recomenda-se então que o médico inicie os atendimentos um pouco depois das chegadas dos clientes, de forma a reduzir essa ociosidade.

Na simulação foi considerado que os pacientes chegavam por volta das 13 horas, que o médico começava os atendimentos às 13:30 e colocou-se a restrição para as chegadas dos pacientes só ocorrerem até as 15:30 horas.

5 CONCLUSÕES

O trabalho mostrou a relevância do estudo de filas em consultórios particulares e os ganhos de desempenho que podem ser obtidos através de uma abordagem adequada, além da mensuração dos atuais problemas e melhor visibilidade da situação.

Para a melhoria do sistema propõe-se então o balanceamento da taxa de chegada através da implementação de um sistema de marcações, ou simplesmente da orientação aos pacientes que ligarem para chegar em horários distribuídos. Dessa forma pode-se evitar a chegada concentrada no horário de início de atendimento do consultório e diminuir a espera e o número de pacientes na fila

Deve-se levar em consideração que o modelo proposto não será válido para os casos em que ocorrerem surtos de determinadas doenças que aumentem substancialmente a demanda, pois a taxa de chegada pode superar a taxa de atendimento.

Foi observado que caso o tempo de atendimento não fosse curto os problemas com espera seriam ainda maiores, pois se provocaria uma instabilidade no sistema. Não foi identificada de forma precisa alguma informação em relação aos pacientes que pudesse permitir uma orientação do atendimento do consultório, embora tenha sido observado que nos casos das quintas-feiras, de pacientes mulheres e de pacientes com acompanhantes o tempo de atendimento possuía uma leve tendência a ser maior.

O presente trabalho trouxe como principal dificuldade os prazos curtos para sua elaboração, interferindo diretamente na quantidade de dados coletados. Sendo assim não foram possíveis considerações mais precisas sobre as características do sistema em estudo e na maioria dos casos observaram-se tendências que poderiam ser atestadas através de um maior volume de dados na coleta.

Houve problemas iniciais no preenchimento de um dos campos do formulário, o referente ao retorno dos pacientes, pois a secretária responsável pela coleta entendeu de forma equivocada a intenção desse campo e, devido ao curto prazo da coleta, não foi possível realizar um novo levantamento de dados que pudessem compensar essa falha. Além disso, ocorreu, em um dia, de a demanda do consultório ser bastante elevada e a coleta de dados ser prejudicada pelo alto movimento, visto que a pessoa responsável pela coleta não o pode fazer de forma correta. Os dados desse dia acabaram por não ser aproveitados.

Propõe-se para sugestões de trabalhos futuros uma pesquisa mais extensa e aprofundada dos dados, de forma a identificar correlações precisas entre eles que permitam uma orientação das marcações das consultas, e a utilização de testes de hipóteses para validação final dos resultados, permitindo a identificação do modelo de filas mais adequado para o caso. Sugere-se também a implementação no consultório das sugestões de melhorias citadas para o monitoramento e avaliação da efetividade destas.

REFERÊNCIAS BIBLIOGRÁFICAS

CAMPOS, M. S. O fantástico box-plot. **Boletim da Siqueira Campos**, ano V, n.8; jan. 2001. Disponível em: “<http://www.siqueiracampos.com/pdf/b8.pdf>”. Acessado em: 26/11/2007

CARVALHO, L. C.; BOFILL, S.; ALMEIDA, T. L. Avaliação de ambientes de pronto-atendimento hospitalar sob a visão dos pacientes. **Série: Textos de Alunos de Psicologia Ambiental**, n.5, Brasília, DF: Universidade de Brasília, 2001.

COSTA, L. C. **Teoria das filas**. São Luis, MA: Universidade Federal do Maranhão. Disponível em: <http://www.decom.ufop.br/prof/rduarte/CIC271/TeoriaFilas_Cajado.pdf>. Acessado em: 02/10/2007.

DOY, F. E.; BRESSAN, G.; PEREIRA, G. H. A.; MAGALHÃES, M. N. **Simulação do serviço de correio eletrônico através de um modelo de filas**; Pesquisa Operacional. vol.26, n.2, p.241-253, Maio/Ago. 2006.

FOGLIATTI, M. C.; MATTOS, N. M. C. **Teoria das filas**. Rio de Janeiro: Interciência, 2007.

JACOB, A. V. et al. Avaliação da satisfação dos clientes e diagnóstico de filas no restaurante universitário noturno da UFSC. **VI SEPROSUL – Semana de Engenharia de Produção Sul-americana**, Florianópolis, SC: UFSC, novembro, 2006. Disponível em: “<http://www.eps.ufsc.br/~peteps/Artigos/Artigo%20RU.pdf>”. Acessado em 02/10/07.

MENDES, A. B. **Investigação operacional: distribuição exponencial**, Universidade dos Açores, 27/04/2006. Disponível em: “www.uac.pt/~amendes/invop/filas_intro.pdf” Acessado em 02/10/2007.

MORABITO, R.; LIMA, F. C. R. de. Um modelo para analisar o problema de filas em caixas de supermercados: um estudo de caso; **Pesquisa Operacional**; vol.20; n.1; p.59-71, Jun. 2000.

MRTVI, V. O. **O impacto da percepção do tempo de espera para atendimento em clínicas médicas na avaliação da qualidade do serviço pelo consumidor.** VI SemeAD. Disponível em: “<http://www.ead.fea.usp.br/Semead/6semead/MKT/059Mkt%20-%20O%20impacto%20da%20Percep%20do%20Tempo.doc>”. Acessado em: 02/10/2007

NIE, W. Waiting: integrating social and psychological perspectives in operations management. **Omega: The International Journal of Management Science.** n.28, p.611-629, 2000.

PINHEIRO, G. **Teoria de filas e sistemas de comunicação,** DETEL – Departamento de Engenharia Eletrônica e Telecomunicações, novembro, 2006. Disponível em: <<http://www.lee.eng.uerj.br/~gil/Filas.pdf>>. Acessado em 02/10/07.

SARMENTO JR, K. M. de A.; TOMITA, S.; KOS, A. O. de A. O problema da fila de espera para cirurgias otorrinolaringológicas em serviços públicos; **Revista Brasileira de Otorrinolaringologia,** vol.71, n.3, parte 1, p.256-262, Maio/Jun. 2005.

SILVA, V. M. C. F. **A relação médico-paciente.** 2001. Disponível em: “<http://www.sab.org.br/med-terap/art-vitor.htm>”. Acessado em: 01/11/2007

APÊNDICE A

Planilha de coleta de dados

DATA _____

ORDEM PACIENTE	SEXO (F/M)	DATA DE NASCIMENTO	CONVENIO (S/N)	HORARIO DE CHEGADA	HORARIO INICIO ATENDIMENTO	HORARIO FIM ATENDIMENTO	RETORNO (S/N)	ACOMPANHANTE (S/N)
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								

APÊNDICE B

Resultados das simulações

Situação inicial: $\lambda_1 = 9$, $d = 0,06$, $\mu = 6,06$

Simulação	Pacientes	Tempo entre chegadas	Atendimento	Espera	Fila	Utilização	Fim do Atendimento
1	20	5,93	6,80	50,43	6,15	100,0%	15,77
2	20	6,18	6,92	21,36	3,35	98,4%	15,85
3	13	9,07	6,98	7,89	1,69	100,0%	15,01
4	30	4,13	7,48	73,12	9,57	100,0%	17,24
5	14	7,81	8,54	24,87	3,43	85,7%	15,81
6	22	4,51	7,27	42,57	6,09	100,0%	16,16
7	18	6,12	8,35	28,41	3,44	100,0%	16,00
8	19	3,43	8,77	70,04	8,26	100,0%	16,28
9	22	5,84	7,78	40,07	5,23	100,0%	16,35
10	17	5,16	8,71	36,25	4,12	100,0%	15,97
11	15	6,08	6,20	12,70	2,20	100,0%	15,05
12	16	6,72	7,24	35,29	4,13	100,0%	15,43
13	18	6,54	6,56	26,27	3,72	100,0%	15,47
14	17	6,93	8,02	30,24	3,88	100,0%	15,77
15	17	7,84	9,32	35,28	3,59	100,0%	16,14
16	22	5,44	6,70	21,60	3,64	99,1%	15,98
17	18	6,33	9,80	40,73	4,94	100,0%	16,44
18	15	6,53	7,09	15,60	2,67	100,0%	15,27
19	18	7,40	7,82	35,70	4,50	100,0%	15,85
20	13	9,83	7,41	12,41	1,23	90,3%	15,32
21	14	8,16	6,02	12,25	2,00	79,6%	15,27
22	19	6,33	7,82	40,29	4,63	100,0%	15,98
23	17	8,05	7,63	22,32	2,76	84,7%	16,11
24	15	7,55	7,61	21,78	2,67	93,8%	15,55
25	16	8,93	8,44	15,06	1,63	85,7%	16,18
26	20	6,20	6,89	25,57	3,80	100,0%	15,80
27	14	9,26	8,04	17,83	1,93	82,8%	15,79
28	17	4,70	5,84	41,35	6,12	100,0%	15,15
29	20	6,18	8,81	54,47	6,15	100,0%	16,44
30	15	8,60	8,31	20,88	2,67	100,0%	15,58
Média	17,7	6,73	7,64	31,09	4,01	96,7%	15,83

Situação proposta: $\lambda_1 = 5,57$, $d = 0$, $\mu = 6,06$

Simulação	Pacientes	Tempo entre chegadas	Atendimento	Espera	Fila	Utilização	Fim do Atendimento
1	16	9,32	7,79	19,95	3,00	98,1%	15,62
2	19	7,65	7,88	23,95	3,32	100,0%	16,00
3	15	10,00	8,40	16,53	1,73	91,3%	15,84
4	13	10,80	5,87	4,24	1,00	69,2%	15,38
5	14	10,30	6,54	27,04	3,07	83,2%	15,41
6	14	10,28	9,58	13,59	1,64	78,8%	16,26
7	14	10,57	8,47	17,07	1,71	98,5%	15,51
8	21	6,84	7,34	24,70	3,48	99,7%	16,08
9	18	8,27	6,32	12,86	1,67	88,3%	15,71
10	12	12,38	10,46	17,84	1,92	100,0%	15,59
11	17	6,90	6,51	17,77	2,00	97,6%	15,40
12	10	14,87	6,36	3,85	0,30	59,2%	15,63
13	13	11,36	6,22	7,06	0,85	71,4%	15,55
14	14	10,67	7,88	24,41	2,93	87,1%	15,65
15	14	9,75	7,91	20,94	2,29	99,4%	15,36
16	17	8,31	7,52	17,68	2,35	93,2%	15,81
17	16	8,43	5,43	4,11	0,88	70,3%	15,65
18	22	6,18	8,26	28,20	3,91	98,7%	16,58
19	15	9,90	7,48	16,41	2,27	91,0%	15,57
20	15	8,91	8,33	12,23	1,27	87,2%	15,94
21	18	8,11	8,49	33,13	3,33	95,6%	16,18
22	14	10,05	7,47	8,44	1,00	80,1%	15,74
23	18	8,12	7,40	20,67	2,11	90,2%	16,00
24	19	7,36	8,75	66,00	6,32	100,0%	16,27
25	17	8,13	9,07	27,29	2,59	87,4%	16,50
26	15	8,87	6,40	10,26	1,20	73,5%	15,82
27	16	8,38	7,98	15,18	1,94	91,7%	15,86
28	17	6,25	11,16	45,72	4,29	100,0%	16,66
29	17	8,22	7,63	21,68	3,59	100,0%	15,66
30	18	7,57	6,30	28,56	4,06	98,4%	15,43
Média	15,93	9,09	7,71	20,25	2,40	89,3%	15,82