

## 2 Medição do Risco

Neste capítulo as definições das variáveis que serão utilizadas na elaboração do método de classificação de risco proposto, serão apresentadas.

Na primeira seção aparecem algumas definições referentes a seguros. Inicialmente serão apresentados aspectos que definem o perfil dos contratos e, em seguida, as variáveis que são utilizadas na medição do risco.

Na seção seguinte serão apresentadas a definição de classificação de risco, e a função da classificação e suas particularidades, sendo a individualização a característica de principal interesse deste estudo.

### 2.1. Definições

O ramo de atividade de seguros apresenta definições que diferem bastante das demais atividades que compõem o setor de serviços. Por esta razão, esta seção apresentará algumas definições, complementares as apresentadas no capítulo 1, que serão amplamente utilizadas no decorrer desta dissertação.

Inicialmente, serão definidas algumas das variáveis que representam o perfil dos contratos, e em seguida serão apresentadas as variáveis utilizadas na medição do risco.

#### 2.1.1. Variáveis de perfil das apólices

As variáveis que definem o perfil das apólices abrangem características do veículo segurado, do tipo de contrato, e do condutor principal do veículo. Embora exista uma infinidade de variáveis disponíveis, somente a definição de 11 delas será apresentada, já que somente estas serão utilizadas.

A variável “*tipos de veículos*” identifica a associação entre veículos semelhantes. Por exemplo, o Gol 1000 Ouro 16 válvulas 4 portas encontra-se no mesmo grupo do Gol 1000 MI 16 válvulas 4 portas geração III.

A variável “*regiões de tarifação*” identifica as regiões de circulação dos veículos. Um dos fatores relevantes na construção desta variável é a proximidade geográfica entre regiões<sup>7</sup>.

A variável “*idade dos veículos*” apresenta, geralmente, 12 níveis, um para os veículos zero quilômetro<sup>8</sup>, e os outros 11 descrevendo as idades dos veículos, de novos usados<sup>9</sup> até 10 anos ou mais.

A variável “*importância segurada*” define o valor de indenização a ser pago ao segurado, em caso de perda total ou roubo do veículo. Esta variável também é conhecida como limite máximo de indenização.

A variável “*forma de indenização*” apresenta uma relação direta com a importância segurada. É ela que identifica se em caso de roubo ou perda total, o segurado será indenizado com base no valor de mercado do veículo, ou através de valor previamente determinado.

A franquia é o valor de participação do segurado em casos de danos parciais. A franquia pode assumir um valor fixo ou um valor percentual do dano, variando de companhia para companhia, e de contrato para contrato. O “*tipo de franquia contratada*”<sup>10</sup> determina se a franquia é obrigatória ou reduzida<sup>11</sup>.

A variável “*tipo de renovação*” determina se o contrato é uma renovação de outra companhia, da própria seguradora, ou um seguro novo.

A variável “*forma de pagamento do seguro*” define em quantas parcelas o prêmio será fracionado. O fracionamento pode ser feito em até 12 parcelas mensais.

A “*classe de bônus*” determina o número total de contratos continuados nos quais não houve sinistros. Em caso de sinistro, o segurado perderia uma classe de bônus. As classes de bônus variam de 1 a 9.

As variáveis “*idade do condutor principal*” e “*sexo do condutor principal*” não exigem maiores explicações.

---

<sup>7</sup> Atualmente, grande parte das seguradoras brasileiras busca identificar as regiões de tarifação através de códigos postais.

<sup>8</sup> Veículos zero quilômetro são aqueles cujo primeiro proprietário é o atual, e que nunca trafegaram por vias públicas, exceto sob a responsabilidade do fabricante e por pequenas distâncias. É comum o mercado estabelecer um limite de quilometragem para diferenciar veículos zero quilômetro de veículos novos usados.

<sup>9</sup> Veículos novos usados são aqueles que atendem a definição de veículos zero quilômetro, porém que transcendem o limite de quilometragem estabelecido pelas seguradoras.

<sup>10</sup> Uma melhor definição desta variável será apresentada no capítulo 4.

<sup>11</sup> Em algumas seguradoras, há ainda a opção de franquia majorada.

### 2.1.2. Variáveis de risco

As variáveis que definem o risco contabilizam o número de sinistros ocorridos e os valores das indenizações associados, o período de exposição ao risco, o prêmio de risco e a taxa de risco.

Das variáveis disponíveis, apenas as referentes à cobertura de casco, serão consideradas.

Define-se “*sinistro*” como o evento responsável por algum dano sobre o bem segurado, ou mesmo qualquer dano provocado ao bem de terceiros pelo veículo segurado.

Este estudo contemplará apenas três tipos de sinistros: sinistros de perda parcial, sinistros de perda total, e sinistros de roubo. Conforme mencionado, os tipos de sinistro referem-se apenas à cobertura de casco, portanto, não serão considerados sinistros de responsabilidade civil. Também não serão analisados sinistros de incêndio e roubo parcial.

O valor percentual dos danos sobre a importância segurada define a diferença entre sinistros de perda parcial e sinistros de perda total. O sinistro de perda parcial é aquele que totaliza danos inferiores a um dado limite percentual, que normalmente varia entre 65% e 70%. Eventos de valor superior a este limite são considerados sinistros de perda total.

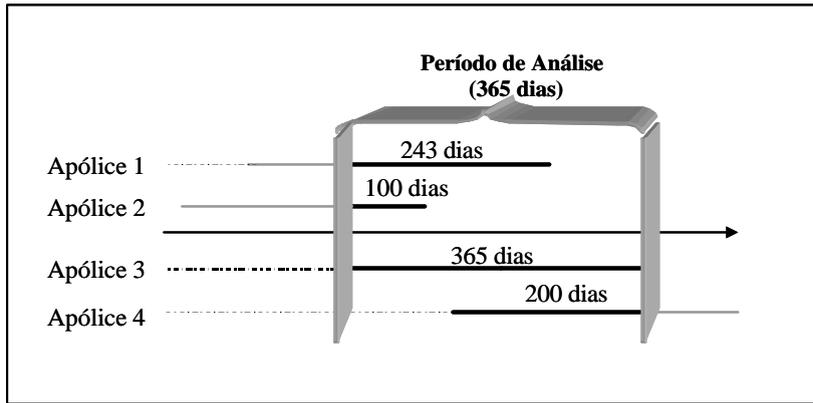
Sinistros de roubo são todos aqueles que resultam da subtração do bem segurado, entendendo por subtração, o roubo ou furto do veículo.

Serão utilizadas três variáveis para determinar o “*número de sinistros*”, cada qual referindo-se a um dos tipos de sinistros apresentados.

Define-se indenização como o ressarcimento, seja monetário ou através de reparos, de danos ocorridos ao veículo segurado. Três variáveis que determinam o “*valor de indenização*” serão consideradas, cada qual associada a um tipo de sinistro específico.

A “*exposição ao risco*” é definida como a proporcionalidade entre período de vigência das apólices e o período de análise. Na figura abaixo, são apresentados exemplos de exposição ao risco.

Figura 1: Exemplos de cálculo da exposição ao risco.



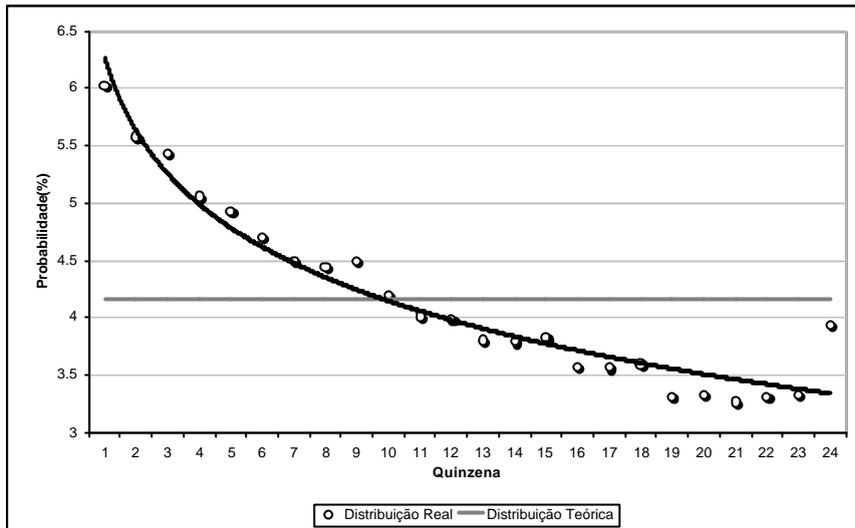
Na figura 1, as linhas contínuas bicolores representam a vigência das apólices. A parte clara das linhas representa o período de vigência ocorrido fora do período de análise e a parte escura define o período de vigência contido no período de análise. O período de análise considerado neste exemplo contempla 365 dias. Desta forma, a exposição ao risco da Apólice 2 seria representada pelos 100 dias de vigência contidos no período de análise, divididos pelos 365 dias que compõe o período, totalizando 0.2739 de exposição ao risco.

Conforme a definição teórica de exposição, supõe-se que a probabilidade de ocorrência de sinistros durante a vigência de uma apólice apresenta uma distribuição uniforme. O que significa que a probabilidade de ocorrência seria a mesma, para cada dia de vigência.

Porém, conforme sugere o gráfico abaixo, a exposição ao risco tende a apresentar distribuição de probabilidade Gama <sup>12</sup>.

<sup>12</sup> Este postulado é demonstrado por Meyer (1983). O autor demonstra que o tempo necessário para se obter um número específico de ocorrências de um evento de Poisson, neste caso, sinistros, apresenta distribuição Gama, sendo a distribuição exponencial um caso particular.

Figura 2: Distribuição de sinistros segundo o período de vigência das apólices.



De acordo com a figura, a probabilidade de ocorrência de sinistros é maior no início de vigência das apólices, decrescendo exponencialmente com o tempo. Além disso, é curioso que ocorra um aumento na probabilidade de sinistros, justamente na última quinzena de vigência das apólices.

Se a distribuição exponencial confirma que os sinistros ocorrem conforme um processo de Poisson, suspeita-se que o comportamento apresentado no final do período esteja relacionado com fraudes, embora não tenham sido encontrados na literatura trabalhos que comprovem esta suposição. De qualquer forma, a hipótese de distribuição uniforme, assumida classicamente, não corresponde ao caráter Poissoniano do processo estocástico implícito. Mesmo assim será utilizada neste trabalho, por simplicidade, porque é a hipótese tradicional, e porque o uso da distribuição exponencial requer um maior desenvolvimento teórico, além de proposta desta dissertação. Em suma, como este estudo não pretende investigar os efeitos do cálculo da exposição sobre o risco, será considerada a definição teórica.

Diante das definições de indenização e exposição, define-se como “*prêmio de risco*” a indenização média de sinistros por unidade exposta. Nesta dissertação serão considerados três tipos de prêmio de risco, cada qual relacionado com os tipos de sinistro mencionados. O prêmio de risco total, representado pela soma

dos prêmios de risco por tipo de sinistro, é parte integrante do prêmio emitido<sup>13</sup>, e é utilizado como forma de cobrar dos segurados o risco inerente ao contrato.

Outra forma de obter o prêmio de risco seria a através da “*taxa de risco*”. A taxa de risco pode ser calculada da seguinte forma: Seja o quociente entre a indenização média por sinistro e a importância segurada ponderada pela exposição. Este quociente é conhecido como grau médio de dano, ou “*severidade*”. Através do produto entre o grau médio de dano e o número de sinistros seria obtido o grau de dano total, que dividido pela exposição, resultaria na taxa de risco. Multiplicando a taxa de risco pela importância segurada, seria obtido o prêmio de risco.

É importante notar que a medição do risco, dada tanto através de taxas quanto através do prêmio de risco, só faz sentido em um dado nível de agregação dos riscos em classes. De outra forma, a própria definição de prêmio/taxa de risco ficaria comprometida. Por definição, os cálculos de prêmio/taxa de risco consideram a exposição no denominador, e quando a exposição assume valor inferior a unidade, representando o caso individual, o valor do prêmio/taxa de risco aumenta, podendo alcançar valores em desacordo com a realidade. Sendo assim, o nível de agregação define as classes de risco, e o conjunto de classes define a estrutura de classificação do risco.

## **2.2. Classificação do risco**

A classificação de risco é a agregação do risco em classes. As classes são geralmente formadas considerando características específicas, que se acredita serem capazes de discernir o maior risco do menor. Portanto, a medição do risco é realizada considerando o comportamento das variáveis de risco, diante da classificação do mesmo.

Não existem regras que direcionem a construção da classificação de risco. A formação das classes reflete apenas os interesses e crenças de cada companhia. Sendo assim, as seguradoras têm liberdade para optar por uma estrutura de classificação de risco mais individualizada, caracterizada pelo grande número de

---

<sup>13</sup> Além da parte referente ao risco, o prêmio emitido leva em consideração agravos referentes às demais despesas da seguradora, e descontos concedidos.

classes; ou optar por uma classificação agregando diferentes tipos de automóveis e situações, diluindo o risco das classes entre os contratos que delas fazem parte.

Visando esclarecer os conceitos de diluição e individualização do risco, serão apresentados dois casos extremos de classificação: a diluição total e a individualização total.

A diluição total do risco pode ser entendida como o caso onde todos os riscos são considerados iguais, e conseqüentemente, todos os segurados pagam taxas idênticas. Neste caso, contratos de menor risco são penalizados e contratos de maior risco; beneficiados. Portanto, este tipo de estrutura é caracterizado pela ausência de parcimônia entre taxas e risco.

A individualização total do risco representa o caso oposto, onde cada contrato é considerado como um risco distinto. Caso fosse possível medir o risco desta forma, prevaleceria a parcimônia entre taxas e risco. Porém, não existem métodos de estimação que, mediante tão pouca informação, resultem em estimativas confiáveis.

Grande parte das seguradoras brasileiras adota uma estrutura de classificação individualizada. Suspeita-se que esta opção seja resultante da aversão ao risco conjugada com estratégias de mercado. Porém, conforme mencionado, esta associação dificulta a medição do risco e, conseqüentemente, a tomada de decisão.

Muitas são as formas de classificação de risco adotadas pelo mercado segurador brasileiro. Classificações formadas pela conjugação de faixas de importância segurada com regiões geográficas, com idade dos veículos, com regiões definidas por código postal, ou mesmo com tipos de veículos, formam algumas das opções mais utilizadas. Porém, visando apenas a dinâmica de preços e a parcimônia entre taxas e risco, nem sempre a atenção devida é dispensada para as conseqüências inerentes à individualização.

Caso não haja informação suficiente por classe, como estimar o risco? Como discernir o risco entre as classes, se existe uma grande semelhança entre elas? Como diferenciar o movimento aleatório do movimento sistemático do risco se, com poucas observações, dificilmente seria possível construir estatísticas confiáveis?

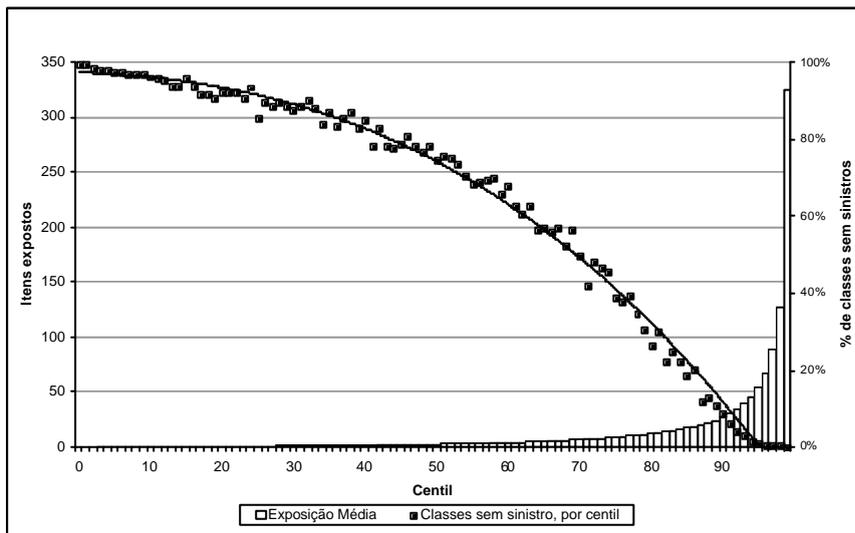
### 2.2.1. Conseqüências da individualização do risco

Supõe-se que a individualização do risco possa ser responsável pela geração de *outliers* e pelo elevado percentual de classes sem sinistros.

É fato que a individualização do risco é responsável pela redução da exposição por classe. Analisando dados do mercado de seguros, verificou-se que 29% das classes apresentaram exposição inferior à unidade. Este fato pode ser interpretado como se as classes representassem apenas o risco individual, ou mesmo como se a probabilidade de ocorrência de sinistros fosse ínfima. Independente da interpretação, a conclusão é única: a estrutura de classificação individualizada produz a redução de informações por classe.

A estratégia utilizada para avaliar a relação entre a distribuição do risco e o número de classes sem sinistros foi separar as classes por centil da exposição, e observar o percentual de classes sem sinistros em cada centil.

Figura 3: Percentual de classes sem sinistros por centil da exposição.



Na figura acima, à esquerda, o eixo vertical define o número médio de itens expostos por centil, representados pelas barras verticais. À direita, o eixo vertical define o percentual de classes sem sinistros, representado pelos pontos e pela curva de tendência. Ainda de acordo com a figura, o eixo horizontal representa os centis de exposição ao risco.

Observando a figura 3, nota-se que 100% das classes de risco que compõe o primeiro centil não apresentam sinistros, e que a exposição média neste centil assume valores aproximadamente nulos. Em oposição, pode ser visto que no último centil a exposição média é de aproximadamente 325 itens, e que praticamente todas as classes de risco que compõe este centil apresentam ao menos um sinistro. Através do comportamento representado pelo gráfico, pode-se suspeitar que haja correlação entre a exposição e o percentual de classes sem sinistros.

Dando continuidade foi avaliada a influência da distribuição da exposição sobre o percentual de *outliers*.

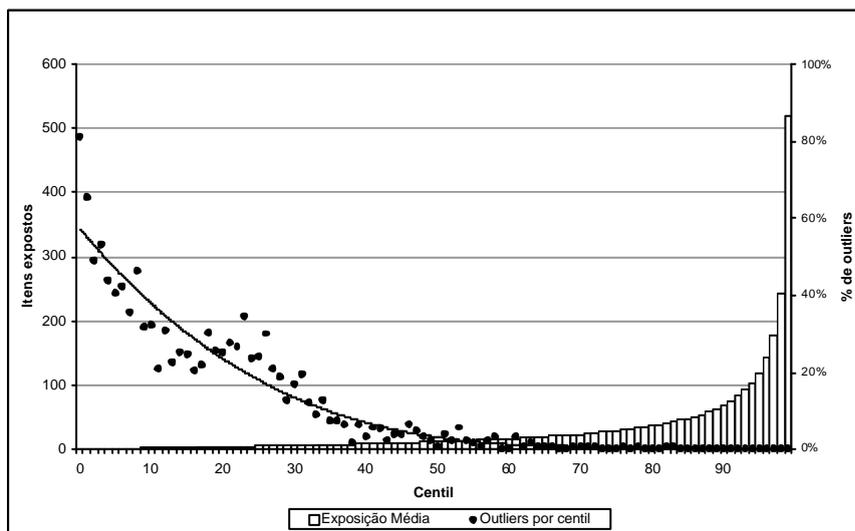
*Outliers* poderiam ser entendidos como valores inesperados causados por problemas de coleta de informações, ou valores exorbitantes e incoerentes com sua distribuição. Neste caso, serão consideradas *outliers* as taxas de risco com valor superior a  $O_3$ , onde  $O_3$  será definido da seguinte forma:

$$O_3 = Q_3 + 1.5(Q_3 - Q_1);$$

onde  $Q_i$  representa o  $i$ -ésimo quartil da distribuição das taxas de risco.

Em seguida, a mesma estratégia de análise adotada na avaliação das classes de risco sem sinistros foi aplicada para avaliar a geração de *outliers*. Porém, as classes sem sinistros tiveram de ser excluídas, caso contrário, devido ao seu grande número, grande parte das classes que apresentassem taxas diferentes de zero seriam consideradas *outliers*.

Figura 4: Percentual de *outliers* por centil da exposição.



Assim como na figura 3, a figura acima apresenta indícios de que o percentual de *outliers* apresenta-se correlacionado com a exposição por classe e, conseqüentemente, com a individualização do risco.

Para corroborar estatisticamente o que pôde ser observado, foram realizadas duas regressões<sup>14</sup>, uma para explicar o percentual de classes sem sinistros através da exposição e outra para explicar o percentual de *outliers* pela exposição.

A tabela abaixo apresenta os coeficientes de determinação obtidos, e outros resultados relevantes.

Tabela 1: Resumo de resultados .

	Classificação completa (1)	Classificação alternativa (2)	Informação perdida de (1) para (2)
Tipos de Veículos	164	158	3.7%
Classes de Risco	34475	12339	64.2%
Exposição (E)	428008	367202	14.2%
Outliers (O)	5641	1443	74.4%
Sem Sinistros (S)	22136	0	100.0%
* R <sup>2</sup> da regressão de O em E	-	<b>0.74</b>	-
* R <sup>2</sup> da regressão de S em E	<b>0.91</b>	-	-

Nota: Em (1) foram consideradas todas as classes que compõem a classificação, e em (2) foram excluídas as classes sem sinistros.

\*Coeficientes de determinação das regressões.

A tabela acima mostra que, caso fossem desconsideradas as classes de risco sem sinistros, 64.2% das classes seriam excluídas da análise e, conseqüentemente, 3.7% dos tipos de veículos. A perda de conhecimento da exposição ao risco representaria 14.2% do total de itens expostos.

A tabela acima reporta ainda que, com a exclusão de classes sem sinistros, a redução no número *outliers* seria de 74.4%.

O valor do coeficiente de determinação obtido através da regressão entre o número de *outliers* e a exposição por classes corrobora a suspeita de correlação. O

<sup>14</sup> Nas regressões foi utilizada a transformação logarítmica das variáveis dependentes. Por este motivo, no caso do percentual de classes sem sinistros foi excluído 1% das observações, e no caso do percentual de *outliers* foram excluídos 27% das observações.

mesmo ocorre quando a regressão considera as classes sem sinistros e a exposição por classe.

Os resultados reportados nesta seção mostram que a individualização do risco resulta na possibilidade remota de ocorrência de sinistros por classe, e por este motivo verifica-se um percentual tão elevado de classes sem sinistros. Além disso, a exposição inferior à unidade, associada à própria definição de taxa e prêmio de risco, apresenta como consequência a geração de *outliers*.

Desta forma, conclui-se que, através da individualização promovida pela classificação original, torna-se inviável determinar o comportamento do risco, já que não há informação suficiente que corrobore qualquer conclusão. O risco observado em dado período provavelmente não apresentará o mesmo comportamento no período seguinte, mesmo que esta classificação seja proposta para uma carteira de risco estável.

### **2.2.2. Estruturas de classificação**

Grande parte das seguradoras brasileiras adota estruturas de classificação de risco semelhantes. Geralmente se utilizam na formação das classes, da combinação das variáveis que identificam os tipos de veículos, as regiões de tarifação e as idades dos veículos. Normalmente o que decorre desta estrutura de classificação é a individualização do risco, ou mesmo problemas de caráter subjetivo. No que tange aos tipos de veículos, o mercado segurador costuma identificá-los ou através de suas características físicas, ou através do valor de mercado dos mesmos.

A variável que identifica os tipos de veículos, conforme descrita na seção 2.1.1, representa um exemplo, já que cria a identificação através da associação de veículos semelhantes.

O que geralmente ocorre quando da construção desta variável é a consideração de detalhes na identificação, o que acaba por reproduzir um número excessivo de tipos, que ao final resulta na individualização.

Quando a identificação de tipos é realizada através do valor de mercado dos veículos, é utilizado como critério a proximidade entre valores médios da importância segurada. Neste caso, não só os perfis de condutores e proprietários

dos veículos representam restrições, como também a utilização dos veículos, o que pode ser entendido como um problema subjetivo de construção.

Quando se trata das regiões de tarifação, normalmente a identificação é formada através da proximidade geográfica entre as áreas. Neste caso a maior restrição está no risco inerente a cada região.

No caso da idade dos veículos, as seguradoras costumam considerar as maiores idades de uma mesma forma, e ainda agrupar o nível intermediário entre zero e 1 ano de idade.

Como pode ser visto, não é comum o uso de técnicas estatísticas aplicadas na formação das variáveis que compõem a classificação de risco.

Diante das possibilidades e das restrições, surgem algumas sugestões de variáveis de classificação, bem como de métodos que poderiam ser utilizados na definição da classificação de risco.

Na identificação de tipos de veículos, o que se pode notar é o excesso de classes devido à pormenorização de detalhes. Uma solução seria agrupá-los, utilizando como variáveis de classificação os atributos que compõe o perfil das apólices. Acredita-se que desta forma seja possível discriminar não só o perfil dos condutores e a utilização dos veículos, mas também o valor de mercado dos mesmos, atendendo assim, não só as expectativas da identificação por semelhança como também as restrições da identificação realizada através do valor de mercado<sup>15</sup>.

Em relação à identificação de regiões de tarifação, surgem duas propostas. Quando as regiões são definidas geograficamente e encontram-se mais agregadas, uma solução seria agrupá-las através de suas frequências de sinistros. Quando as regiões são definidas através de códigos postais, portanto com maior grau de desagregação, o agrupamento poderia ser realizado através da utilização de indicadores de desigualdade social, do perfil da população, de indicadores econômicos e de indicadores de conflitos urbanos. Neste caso, seria imprescindível a utilização de *surveys*<sup>16</sup>, bem como acesso a informações contidas em bancos de dados públicos.

<sup>15</sup> Supõe-se que esta técnica seja mais eficiente quando aplicada por empresas que utilizam canais de vendas bancários, já que estas geralmente apresentam um cadastro de informações sobre seu comportamento financeiro real ao longo do tempo.

<sup>16</sup> Como exemplo de pesquisas nacionais pode ser citado o Censo Demográfico.

No que tange ao agrupamento da idade dos veículos, a proposta é utilizar a depreciação média dos mesmos, ao longo dos anos. Através da depreciação, acredita-se ser possível identificar o grau de semelhança entre as idades.