

UNIVERSIDADE ESTÁCIO DE SÁ

JONIAS DOS SANTOS BUENO

Indicadores Antecedentes na Área de Acidentes de Trabalho na Previdência Social
Brasileira

Rio de Janeiro
2008

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

JONIAS DOS SANTOS BUENO

Indicadores Antecedentes na Área de Acidentes de Trabalho na Previdência Social
Brasileira

Dissertação apresentada à Universidade Estácio de Sá,
como requisito parcial para a obtenção do grau de Mestre
em Administração e Desenvolvimento Empresarial.
Orientador Prof. Dr. Antônio Carlos Magalhães da Silva

Rio de Janeiro
2008



UNIVERSIDADE
Estácio de Sá

PROGRAMA DE PÓS-GRADUAÇÃO EM
ADMINISTRAÇÃO E DESENVOLVIMENTO EMPRESARIAL

A dissertação

**INDICADORES ANTECEDENTES NA ÁREA DE
ACIDENTES DE TRABALHO NA PREVIDÊNCIA SOCIAL BRASILEIRA**

elaborada por

JONIAS DOS SANTOS BUENO

e aprovada por todos os membros da Banca Examinadora foi aceita pelo Curso de Mestrado Profissional em Administração e Desenvolvimento Empresarial como requisito parcial à obtenção do título de

MESTRE EM ADMINISTRAÇÃO E DESENVOLVIMENTO EMPRESARIAL

Rio de Janeiro, 03 de dezembro de 2008.

BANCA EXAMINADORA

Prof. Dr. Antonio Carlos Magalhães da Silva
Presidente
Universidade Estácio de Sá

Prof. Dr. Jesús Domach Moré
Universidade Estácio de Sá

Prof. Dr. Osmani Teixeira de Carvalho Guillén
Unifac/RJ

Dedicatória

Aos meus filhos,
Débora e Diogo,
e à minha esposa
Rosana Gauziski,
partes inseparáveis da minha vida.

AGRADECIMENTOS

Ao eterno amigo João Carlos Beltrão da Costa, que já não se encontra entre nós, pelo apoio, entusiasmo e ênfase na realização deste estudo.

Ao orientador e amigo Antônio Carlos Magalhães da Silva, pelo sólido suporte didático e pela orientação firme e segura na condução desta pesquisa.

A José Augusto Vieira (DATAPREV) pelas valiosas diretrizes na fundamentação e condução do estudo.

À Professora Vera Maria Medina Simonetti pelo apoio na revisão metodológica e pela enorme dedicação, aspectos fundamentais e motivadores na realização deste estudo.

Aos Professores Harvey e Ribas pelas valiosas sugestões e propostas sobre modelos inferenciais multivariados.

A Pedro Márcio Novaes de Moraes (DATAPREV) pela confiança e pelo enorme apoio durante a elaboração deste estudo.

A Jaime Pereira Ferreira (DATAPREV) pelo apoio e pelas intervenções decisivas no embasamento e auxílio na realização do trabalho.

A Sérgio Luiz de Melo Monteiro (DATAPREV) pela consultoria decisiva e apoio intermináveis sobre legislação Previdenciária e sistemas de Tecnologia da Informação utilizados pela Previdência Social.

A Rodrigo Novais Coutinho, Diretor de Negócios da DATAPREV, pelo apoio e confiança dispensados, ajudas imperiosas na realização do estudo.

A Benedito Adalberto Brunca, Diretor do INSS, pelo acesso às informações Previdenciárias, subsídios essenciais para a realização deste estudo.

Ao Coordenador Geral do NMG – Novo Modelo de Gestão, Alan Nascimento da Silva pelo apoio nas análises sobre a Previdência Social.

A Jorge Correa de Paula (DATAPREV) e Luiz Antônio Basílio da Silva (DATAPREV), pelos valiosos auxílios e sugestões na editoração eletrônica.

A Maria Francinete Batista Damião (INSS), pelo auxílio sobre a gestão dos Benefícios por Incapacidade na Previdência Social Brasileira.

A Wagner de Lemos Liebert Júnior (DATAPREV) e a Valéria Blini (DATAPREV), pelo apoio na obtenção e validação das informações que embasam este estudo.

Aos grandes amigos Hélcio Ribeiro Rosa e Rosângela Monlevade Silva pelo apoio e pelo inabalável entusiasmo durante a condução deste estudo.

EPÍGRAFE

O acaso não existe.
(Alan Kardec)

Resumo

O propósito deste estudo é avaliar a similaridade entre os benefícios concedidos pela Previdência Social e os cessados devido às irregularidades identificadas pela Auditoria ou pela Inspeção do Instituto Nacional do Seguro Social ao longo dos anos, comparando-os com os benefícios válidos e vigentes na categoria de benefício por incapacidade. Como resultado desta análise, foi gerado um modelo de Regressão Logística capaz de indicar o grau de predição (*score*), auxiliando os órgãos de controle da Previdência Social na seleção da amostragem de benefícios em trabalhos futuros. O modelo resultante, foi aplicado sobre os benefícios mantidos pelo INSS com os propósitos de verificar o percentual identificado e os valores que poderiam ser ressarcidos ou não-pagos, no caso de uma análise individual. Foram avaliados todos os benefícios por incapacidade concedidos pela Previdência Social no período de junho/1994 até jan/2008, sendo o modelo adequado somente a quatro modalidades de benefícios por incapacidade. O modelo visa a geração de escalas (*scores*) de risco para os benefícios, possibilitando ações mais pontuais dos órgãos de controle e Auditoria do Ministério da Previdência e Assistência Social. O modelo utilizou 20 variáveis explicativas indicadas por especialistas da Previdência Social, a partir de um *survey de experiência*. Foram consideradas 10 variáveis explicativas, o nível de predição foi de 96,7% e foram selecionados 65.784 benefícios com *scores* entre 50% e 100%, a partir de um universo de 1,42 milhões de benefícios ativos no período, totalizando R\$ 1,335 bilhões em créditos gerados para os benefícios por incapacidade selecionados.

Palavras-chave: Benefícios por incapacidade; Acidentes de trabalho; Previdência Social; Regressão logística; Amostragem probabilística sistemática.

Abstract

The aim of this study is to evaluate the similarity among the benefits granted by Brazilian Social Welfare and interrupted them due to the identified irregularities for the Auditing or for Inspection of the Brazilian Social Welfare Institute along the years, comparing them with the valid and effective benefits in the category for incapacity. As a result of this analysis, a model of Logistic Regression was generated and this model is capable to indicate the prediction degree (score), aiding the organs of control of Social Welfare in the selection of the sampling of benefits in future works. The resulting model, was applied on the benefits maintained by Social Security Department with the purposes of verifying the percentile identified and the values that could be compensated or unpaid, in the case of an individual analysis. They were appraised all of the benefits for incapacity granted by Social Welfare in the period june/1994 to jan/2008, being the appropriate model only to four modalities of benefits for incapacity. The model seeks the generation of scales (scores) of risk for the benefits, making possible more punctual actions of the Control organs and Auditing of the Ministry of the Social Security. The model used 20 suitable explanatory variables for specialists of Social welfare, starting from a survey of experience. 10 explanatory variables were considered, the prediction level was of 96,7% and 65.784 benefits were selected with scores between 50% and 100%, starting from an universe of 1,42 million active benefits in the period, totaling R\$ 1,335 billion in credits generated for the benefits by incapacity selected.

Keywords: Benefits by incapacity; Work accidents; Social Security; logistic regression; Sampling systematic.

Lista de Tabelas

TABELA 1 - BENEFÍCIOS MANTIDOS PELA PREVIDÊNCIA EM DEZEMBRO/2007	21
TABELA 2 - BENEFÍCIOS POR INCAPACIDADE CONCEDIDOS EM DEZEMBRO/2007	21
TABELA 3 - DEFINIÇÃO INICIAL DAS VARIÁVEIS UTILIZADAS NO ESTUDO	27
TABELA 4 - DEFINIÇÃO DA ORIGEM DAS VARIÁVEIS.....	72
TABELA 5 – RESULTADO DA APURAÇÃO DOS QUESTIONÁRIOS	78
TABELA 6 - DISTRIBUIÇÃO DA POPULAÇÃO PELAS UNIDADES TERRITORIAIS	79
TABELA 7 - POPULAÇÃO OCUPADA POR UNIDADES TERRITORIAIS	80
TABELA 8 - POSIÇÃO DOS BENEFÍCIOS POR INCAPACIDADE	81
TABELA 9 - DISTRIBUIÇÃO DOS BENEFÍCIOS POR INCAPACIDADE POR ESPÉCIE	82
TABELA 10 - BENEFÍCIOS POR INCAPACIDADE CESSADOS DISTRIBUÍDOS POR ESPÉCIE	83
TABELA 11 - BENEFÍCIOS POR INCAPACIDADE ATIVOS POR ESPÉCIE	84
TABELA 12 - BENEFÍCIOS POR INCAPACIDADE CESSADOS APÓS JUNHO/1994	85
TABELA 13 - BENEFÍCIOS POR INCAPACIDADE ATIVOS INICIADOS APÓS JUNHO/1994	86
TABELA 14 - BENEFÍCIOS ATIVOS INTEGRANTES DA AMOSTRA SISTEMÁTICA.....	88
TABELA 15 - COMPOSIÇÃO DA AMOSTRA FINAL DA PESQUISA.....	88
TABELA 16 - MOTIVOS DE CESSAÇÃO DOS BENEFÍCIOS POR INCAPACIDADE	89
TABELA 17 - DISTRIBUIÇÃO DOS BENEFÍCIOS POR ESPÉCIE.....	91
TABELA 18 - DISTRIBUIÇÃO DOS SEGURADOS POR SEXO INFORMADO	91
TABELA 19 - TESTE DE NORMALIDADE KOLMOGOROV-SMIRNOV	113
TABELA 20 - SUMÁRIO DO PROCESSAMENTO DO MODELO.....	114
TABELA 21 - TABELA DE CLASSIFICAÇÃO SOMENTE COM A CONSTANTE.....	115
TABELA 22 - VARIÁVEIS NA EQUAÇÃO COM A CONSTANTE.....	115
TABELA 23 - TESTE DE AVALIAÇÃO SEQUENCIAL DOS COEFICIENTES.....	117
TABELA 24 - RESUMO DO MODELO.....	118
TABELA 25 - RESULTADO DO TESTE DE HOSMER E LEMESHOW	119
TABELA 26 – TABELA DE CONTINGÊNCIA PARA O TESTE HOSMER E LEMESHOW	120
TABELA 27 - CLASSIFICAÇÃO FINAL DO MODELO.....	120
TABELA 28 - VARIÁVEIS INDEPENDENTES INCORPORADAS AO MODELO	124
TABELA 29 - DISTRIBUIÇÃO DOS BENEFÍCIOS POR ESPÉCIE – SCORE ENTRE 50% E 100%.....	127
TABELA 30 -DISTRIBUIÇÃO DOS BENEFÍCIOS POR ESPÉCIE - SCORE ENTRE 90% E 100%.....	127
TABELA 31 - DISTRIBUIÇÃO DOS BENEFÍCIOS POR ESPÉCIE SCORE ENTRE 99% E 100%.....	128
TABELA 32 - DISTRIBUIÇÃO POR SISTEMA DE CONCESSÃO – SCORE ENTRE 90% E 100%	128
TABELA 33 - DISTRIBUIÇÃO DOS BENEFÍCIOS POR TEMPO DE RECEBIMENTO	129
TABELA 34 - BENEFÍCIOS POR FORMA DE FILIAÇÃO - SCORE ENTRE 90% E 100%.....	130
TABELA 35 - DISTRIBUIÇÃO POR RAMO DE ATIVIDADE – SCORE ENTRE 90% E 100%.....	130
TABELA 36 - DISTRIBUIÇÃO POR ACESSO AO CNIS – SCORE ENTRE 90% E 100%.....	131
TABELA 37 – DISTRIBUIÇÃO POR REGIÃO GEOGRÁFICA – SCORE ENTRE 90% E 100%.....	131
TABELA 38 - DISTRIBUIÇÃO DOS SEGURADOS IDENTIFICADOS POR FAIXA DE RENDA	132
TABELA 39 - DISTRIBUIÇÃO DOS SEGURADOS POR VALOR DE CONCESSÃO.....	133
TABELA 40 -DISTRIBUIÇÃO PELA QUANTIDADE DE REVISÕES – SCORE ENTRE 90% E 100%	133

LISTA DE QUADROS

QUADRO 1 - DEMONSTRATIVO DO RESULTADO PREVIDENCIÁRIO.....	19
QUADRO 2 - RESUMO TEÓRICO DO ESTUDO.....	69
QUADRO 3 - VOLUME DE INFORMAÇÕES DISPONÍVEL.....	95
QUADRO 4 - DISTRIBUIÇÃO DOS BENEFÍCIOS POR INCAPACIDADE	96
QUADRO 5 - ANÁLISE DE MULTICOLINEARIDADE DA VARIÁVEIS INDEPENDENTES	111
QUADRO 6 - VARIÁVEIS MÉTRICAS APÓS ANÁLISE DE CORRELAÇÃO (MULTICOLINEARIDADE)	112
QUADRO 7 - VARIÁVEIS CONSIDERADAS NO MODELO FINAL	134

LISTA DE FIGURAS

FIGURA 1 - BENEFÍCIOS CONCEDIDOS PELA PREVIDÊNCIA	20
FIGURA 2 - DEMONSTRAÇÃO DOS ESTÁGIOS DA ANÁLISE PROJETIVA	38
FIGURA 3 - RAZÃO DE CHANCE DA REGRESSÃO LOGÍSTICA	102
FIGURA 4 - DEMONSTRAÇÃO DO MODELO QUI-QUADRADO	106
FIGURA 5 - HISTOGRAMA DOS BENEFÍCIOS COM SCORE ENTRE 50% E 100%.....	125
FIGURA 6 - DISTRIBUIÇÃO DO BENEFÍCIOS NA FAIXA DE RISCO ENTRE 90% E 100%	126
FIGURA 7 - QUANTIDADE DE BENEFÍCIOS POR ANO DE RECEBIMENTO	129
FIGURA 8 - DISTRIBUIÇÃO DOS BENEFÍCIOS PELO VALOR DA RENDA MENSAL.....	132
FIGURA 9 - DISTRIBUIÇÃO DOS BENEFÍCIOS POR QUANTIDADE DE SALÁRIOS-MÍNIMOS	133

LISTA DE SIGLAS

CBO - Cadastro Brasileiro de Ocupações

CNAE – Cadastro Nacional de Atividades Econômicas

CRP – Centro de Reabilitação Profissional

DATAPREV – Empresa de Tecnologia e Informações da Previdência Social

DOU – Diário Oficial da União

IBGE- Instituto Brasileiro de Geografia e Estatística

INPC – Índice Nacional de Preços ao Consumidor

INSS – Instituto Nacional do Seguro Social

MPAS – Ministério da Previdência e Assistência Social

NMG – Novo Modelo de Gestão

SUB – Sistema Único de Benefícios

SUMÁRIO

1	INTRODUÇÃO	14
1.1	Contextualização.....	14
1.2	Problema.....	16
1.3	Objetivos.....	17
1.4	Relevância e justificativa	18
1.5	Hipóteses	24
1.6	Definição das variáveis	26
2	REFERENCIAL TEÓRICO.....	32
2.1	Análise projetiva	32
2.2	Indicadores Antecedentes nas diversas áreas	39
2.3	Indicadores Antecedentes na área social	58
3	METODOLOGIA.....	62
3.1	Método Indutivo	68
3.2	Survey de experiência	70
3.3	Definição das variáveis	71
3.4	Definição do questionário.....	73
3.5	Processamento dos questionários	77
3.6	Pesquisa exploratória.....	79
3.7	Regressão Logística.....	97
4	APLICAÇÃO PRÁTICA	108
4.1	Apresentação do modelo gerado	108
4.2	Aplicação do modelo gerado.....	125
5	CONCLUSÃO.....	135
5.1	Análise dos resultados.....	135

5.2	Considerações Finais	139
6	SUGESTÕES E RECOMENDAÇÕES	141
7	BIBLIOGRAFIA.....	142
8	ANEXOS.....	149
8.1	ANEXO A – Auxílio doença Previdenciário	149
8.2	ANEXO B – Aposentadoria por invalidez acidente do trabalho	151
8.3	ANEXO C - Auxílio por Acidente de Trabalho	152
9	APÊNDICES.....	154
9.1	APÊNDICE A - Questionário.....	154
9.2	APÊNDICE B – Variáveis categóricas do questionário	159
9.3	APÊNDICE C – Processamento dos questionários	164
9.4	APÊNDICE D – Variáveis categóricas agrupadas.....	169
9.5	APÊNDICE E – Análise dos resultados entre 50 e 100%.....	170
9.6	APÊNDICE F – Análise dos resultados entre 99 e 100%.....	174

1 INTRODUÇÃO

1.1 Contextualização

O objetivo deste estudo foi desenvolver um modelo capaz de avaliar a conformidade dos benefícios concedidos e mantidos pela Previdência Social Brasileira na categoria de Benefícios por Incapacidade, a partir da identificação das dimensões comuns em benefícios cessados pelos órgãos de Auditoria do INSS (Instituto Nacional do Seguro Social) e DATAPREV (Empresa de Tecnologia e Informação da Previdência Social). Foi analisada a exatidão das informações prestadas no momento da concessão, no intuito de contribuir para a redução das concessões indevidas dos benefícios por incapacidade.

Este estudo assumiu como corretos os critérios utilizados pela Auditoria do INSS nos processos de análise dos benefícios. Os benefícios foram cessados já que os mesmos não atendiam aos critérios vigentes à época do exame, tendo sido identificadas características que não garantiam o nível de conformidade legal.

A Previdência Social Brasileira foi instituída através do DECRETO No. 4.682 de 24 de Janeiro de 1923 – Publicado no DOU de 28/01/1923, a *Lei Eloi Chaves*, contemplando inicialmente todos os trabalhadores detentores de vínculos empregatícios formais.

O Instituto Nacional do Seguro Social (INSS) e o Ministério da Previdência e Assistência Social (MPAS), foram criados através da Lei nº 8.213/91, do Decreto nº 2.172/97, da Lei nº 8.212/91, e do Decreto nº 2.173/97, que segundo BRASIL (2002) possui a finalidade de:

“Garantir aos seus segurados, sob forma de benefícios e serviços, os meios indispensáveis a sua manutenção, quando na inatividade, por motivo de velhice, incapacidade para o trabalho, ou após o cumprimento do tempo de serviço estabelecido em lei”.

Dentre os diversos tipos de benefícios assegurados aos seus participantes e dependentes estão as aposentadorias, pensões e os Auxílios Assistenciais, distribuídos por áreas de atuação em relação ao local onde as atividades são desenvolvidas pelos empregados e empregadores, sendo considerados Segurados no âmbito da Previdência Social (Área Urbana e Área Rural) os trabalhadores que desempenham atividades remuneradas e que são classificados entre os 04 setores macroeconômicos (*agricultura, indústria, comércio e serviço*). Em suas definições iniciais, a Previdência Social Brasileira consistia de um *modelo de repartição simples*, onde a soma de todos os valores arrecadados destinava-se ao pagamento de benefícios aos seus participantes. Este pressuposto foi embasado por valores atuariais e demográficos vigentes no momento de sua instituição, considerando também indicadores sociais como expectativa de vida, idade inicial de ingresso em atividades laborais, aposentadorias e outros benefícios baseados em tempo de contribuição pelo segurado, nível de risco das atividades desempenhadas pela população e desempenho da economia (nível de desemprego em momentos de retração econômica). As condições de trabalho possibilitavam aos contribuintes, em sua maioria, a solicitação de aposentadorias por tempo de contribuição ou aos seus dependentes a concessão de pensões sobre a aposentadoria inicialmente concedida por tempo de contribuição.

Segundo BRASIL (2002), o modelo previsto de Previdência e Amparo Social, em sua formulação, não considerou alterações sócio-econômicas tais como a relação entre contribuintes e beneficiários que em 1950 era de oito contribuintes

para um beneficiário e em 1997 era de 1,7 contribuintes para um segurado; o aumento da expectativa de vida dos Brasileiros; a extensão da abrangência da Previdência Social para cidadãos que não realizaram contribuições (Aposentadorias de trabalhadores Rurais e Amparos Assistenciais), além da evolução verificada nos últimos 10 anos sobre os benefícios relativos às Aposentadorias por Invalidez e Auxílios por acidente de trabalho.

Os benefícios acidentários ainda são concedidos sem que o contribuinte tenha atingido o volume previsto de contribuições para a solicitação do Benefício por tempo de contribuição, uma vez que a sua capacidade laborativa foi reduzida ou cessada em função da ocorrência de acidente no desempenho da atividade ou devido às doenças decorrentes do desempenho de suas atividades ou em decorrência de fatores externos que impeçam a execução de atividades profissionais.

1.2 Problema

Segundo BOAVENTURA (2007), o ponto de partida para a pesquisa deve ser desdobrado em questões ou hipóteses, fundamentadas na literatura e operacionalizadas pela metodologia e claramente definidos os objetivos a serem alcançados, através do desdobramento em questões que possibilitem a sua compreensão e funcionalidade, permitindo a utilização de instrumentos, técnicas e de processos de investigação. Para Rudio (2001) apud Matias-Pereira (2007), a definição do problema é um momento especial no trabalho do pesquisador, uma vez que a formulação do problema consiste em informar com clareza quais são as dificuldades em questão e as formas pretendidas de resolução, através da limitação do seu campo e apresentação das suas características. Para Matias-Pereira (2007),

a definição do problema deve explicitar quais são as dificuldades em questão e como pretende-se resolvê-las através de uma pesquisa, limitando o seu campo e apresentando as suas características.

Dentro deste contexto, este estudo pretende investigar:

- i) Quais os impactos da aplicação de um modelo estatístico, segundo a Análise Projetiva, na detecção de características comuns entre os benefícios por incapacidade cessados e os mantidos pela Previdência Social Brasileira.

1.3 Objetivos

1.3.1 Objetivos Gerais

- i) Apresentar um mapeamento das variáveis que influenciam na conformidade dos benefícios concedidos e mantidos pela Previdência Social Brasileira nas categorias de benefícios por incapacidade, particularmente nas subcategorias de Auxílio por Acidente do trabalho e Aposentadoria por Invalidez.
- ii) Gerar um modelo estatístico que permita, através da utilização de dados oriundos da Previdência Social, avaliar em profundidade a ocorrência de Acidentes do Trabalho e conseqüentes concessões de Aposentadorias por Invalidez (Benefícios por Incapacidade).

1.3.2 Objetivos específicos

- i) Apontar as variáveis que na opinião dos especialistas, consultados através de questionário, possam influenciar na categorização de um benefício por incapacidade como em conformidade ou não, em relação à legislação vigente;

- ii) Verificar o nível de influência do modelo estatístico gerado e da representatividade estatística das variáveis na identificação de irregularidades quando aplicado aos benefícios concedidos e mantidos;
- iii) Estabelecer *scores* de risco para os benefícios por incapacidade;
- iv) Auxiliar os processos de seleção e análise efetuados pelos órgãos de controle do INSS e DATAPREV;
- v) Promover melhorias no exame de conformidade legal das novas solicitações de benefícios por incapacidade e nos já existentes sob a gestão do INSS;

1.4 Relevância e justificativa

A partir do quadro 1, observa-se o fluxo de caixa da Previdência Social Brasileira em valores nominais. Na coluna (Dez/2007) é apresentado o volume nominal arrecadado oriundo das fontes que compõem as receitas da Previdência Social. Este valor foi inferior ao volume de recursos necessários para efetuar as obrigações Previdenciárias, constante no ítem 2 do mesmo quadro.

O fato descrito acima acarreta em um déficit de R\$ 3.912,70 milhões de Reais no mês de Dezembro de 2007.

O pagamento dos benefícios por incapacidade, contribui significativamente para o déficit previdenciário, uma vez que estes benefícios são pagos sem que as contribuições previstas, pelo prazo definido, tenham sido arrecadadas.

O déficit poderá ser ampliado, à medida que um auxílio por acidente (afastamento por período necessário para recuperação do segurado) transformar-se em uma aposentadoria por invalidez (através de parecer médico, o segurado torna-se inválido para exercer atividades profissionais) ou pensão por morte acidentária

(situação em que o segurado falece devido à ocorrência de acidente ou moléstia decorrente do exercício da atividade profissional).

Quadro 1 - Demonstrativo do Resultado Previdenciário

Resultado do Regime geral de Previdência Social - Em R\$ milhões de Janeiro/2008			
Arrecadação líquida, Benefícios Previdenciários e Resultado Previdenciário do RGPS			
	Jan/2007	Dez/2007	Jan/2008
1 – Arrecadação líquida (1.1 + 1.2 + 1.3 + 1.4)	9.987,5	19.965,1	11.206,9
1.1 – Receitas correntes	10.872,5	20.347,4	12.701,5
1.2 – Recuperação de créditos	623,1	849,4	613,0
1.3 – Restituições de contribuintes	(13,2)	(18,6)	(2,0)
1.4 – Transferências a terceiros	(1.496,7)	(1.213,1)	(2.105,6)
2 – Despesas com benefícios previdenciários (2.1 + 2.2 + 2.3)	13.880,0	23.877,8	16.295,8
2.1 – Pagos pelo INSS	13.703,9	23.666,2	13.857,8
2.2 – Sentenças Judiciais – Justiça Federal	176,1	195,6	2.437,8
2.3 – Sentenças Judiciais – Justiça Estadual	0	16,0	0,2
3 – Resultado Previdenciário (1 – 2)	(3.894,3)	(3.912,7)	(5.088,9)
4 – Despesas com benefícios Assistenciais e EPU	1.179,9	1.911,2	1.263,7

Obs. O item 4 (Despesas com benefícios Assistenciais e EPU) não entra no cálculo do resultado Previdenciário, pois estes benefícios, embora operacionalizados pelo INSS, são pagos com recursos do Tesouro Nacional.

Fonte: INSS (Fluxo de caixa ajustado do Sistema Informar)

Este estudo avaliou as espécies que constituem o conjunto de *Benefícios por Incapacidade*, identificados na figura 1 como Auxílio-doença Previdenciários e os benefícios categorizados como Acidentários, retratando as situações nas quais o segurado é impedido de desempenhar as atividades no trabalho devido a fatores que reduziram as suas funções motoras ou psíquicas, decorrentes de moléstias desenvolvidas no exercício de suas funções ou de acidentes em trânsito ou no próprio local de trabalho, buscando identificar as características dos benefícios cessados pelos Órgãos de Auditoria e Inspeção do INSS, devido à existência de possíveis irregularidades nos mesmos, fato que poderá contribuir para a realização de pagamentos indevidos.

Figura 1- Benefícios concedidos pela Previdência

 Quantidade de Benefícios Concedidos – RGPS – Janeiro/2007, Dezembro/2007 e Janeiro/2008 –					
	jan/07 (A)	dez/07 (B)	jan/08 (C)	Var. % (C/B)	Var. % (C/A)
TOTAL	315.959	306.975	326.902	6,5	3,5
PREVIDENCIÁRIOS + ACIDENTÁRIOS	295.449	280.136	299.852	7,0	1,5
Previdenciários	283.961	257.513	273.714	6,3	(3,6)
Aposentadorias	56.646	74.762	70.982	(5,1)	25,3
Idade	32.707	39.445	35.641	(9,6)	9,0
Invalidez	8.975	15.447	16.992	10,0	89,3
Tempo de Contribuição	14.964	19.870	18.349	(7,7)	22,6
Pensão por Morte	27.182	25.698	28.767	11,9	5,8
Auxílio-Doença	170.110	122.513	140.394	14,6	(17,5)
Salário-Maternidade	28.884	33.333	32.206	(3,4)	11,5
Outros	1.139	1.207	1.365	13,1	19,8
Acidentários	11.488	22.623	26.138	15,5	127,5
Aposentadorias	284	547	533	(2,6)	87,7
Pensão por Morte	103	92	89	(3,3)	(13,6)
Auxílio-Doença	10.363	21.202	24.795	16,9	139,3
Auxílio-Acidente	732	773	710	(8,2)	(3,0)
Auxílio-Suplementar	6	9	11	22,2	83,3
ASSISTENCIAIS + EPU	20.510	26.839	27.050	0,8	31,9
Assistenciais	20.506	26.807	27.000	0,7	31,7
Amparos Assistenciais - LOAS	20.468	26.783	26.958	0,7	31,7
Idoso	11.752	14.092	13.992	(0,7)	19,1
Portador de Deficiência	8.716	12.691	12.966	2,2	48,8
Pensões Mensais Vitalícias	37	23	41	78,3	10,8
Rendas Mensais Vitalícias	1	1	1	0,0	0,0
Idade	-	-	-	-	-
Invalidez	1	1	1	0,0	0,0
Encargos Previdenciários da União (EPU)	4	32	50	56,3	1150,0

Fontes: Anuário Estatístico da Previdência Social - AEPS; Boletim Estatístico da Previdência Social - BEPS.
Elaboração: SPS/MPS.

Fonte: DATAPREV – Sistema Síntese

Em função do conjunto de fatores que contribuem para a concessão e manutenção dos benefícios previdenciários, tais como aspectos legais e mobilidades sócio-econômicas e demográficas, este estudo buscará a análise conjunta e simultânea dos fatores envolvidos, com o propósito de estabelecer relações e influências no comportamento dos benefícios acidentários. A consideração individual dos fatores pelos especialistas, devido à complexidade e ao volume dos fatores, pode ter apresentado baixa eficiência na análise dos benefícios e seus comportamentos. Segundo ZADEH (1973):

À medida que a complexidade de um sistema aumenta, nossa habilidade para fazer afirmações precisas e que sejam significativas acerca deste sistema diminui até que um limiar é atingido além do qual precisão e significância (ou relevância) tornam-se quase que características mutuamente exclusivas.

Segundo o Boletim Estatístico do Ministério da Previdência Social (<http://www.inss.gov.br>), no mês de Dezembro de 2007, foram gerados créditos para 4.261.918 beneficiários relativos à categoria de Benefícios por Incapacidade, totalizando R\$ 2,4 bilhões correspondendo este valor a 19,02% do total de R\$ 12,6 bilhões pagos aos 25.170.283 de benefícios mantidos pela Previdência Social na mesma data.

A composição dos benefícios, em função das categorias que foram alvo do estudo em questão, encontram-se detalhadas na tabela 1.

Tabela 1 - Benefícios mantidos pela Previdência em Dezembro/2007

Tipos de benefícios	Quantidade de créditos	Valores em R\$
31 - AUXILIO DOENCA PREVIDENCIARIO	1.244.350	946.545.677,59
32 - APOSENTADORIA INVALIDEZ PREVIDENCIARIA	2.731.144	1.286.465.676,39
91 - AUXILIO DOENCA POR ACIDENTE DO TRABALHO	137.723	75.744.095,35
92 - APOSENT. INVALIDEZ ACIDENTE TRABALHO	148.701	94.063.567,32
Total de créditos alvo do estudo	4.261.918	2.402.819.016,65
Total de créditos gerados	25.170.283	12.635.504.460,13

Fonte: Boletim Estatístico do Ministério da Previdência Social

Ainda na competência Dezembro/2007 foram concedidos 159.809 Benefícios por Incapacidade, correspondendo a 57,05% do total de novos benefícios na competência. Estes quantitativos referem-se à soma das coberturas Previdenciária e Acidentária, segundo a tabela 2.

Tabela 2 - Benefícios por incapacidade concedidos em Dezembro/2007

Benefícios por incapacidade concedidos em Dezembro/2007	
Benefício por Incapacidade	Quantidade
Aposentadorias por Invalidez Previdenciária	15.547
Auxílio-doença Previdenciário	122.513
Aposentadorias Acidentárias	547
Auxílio-doença Acidentário	21.202
Total	159.809
Total de Benefícios concedidos (Previdenciários + Acidentários)	280.136
Representação do alvo do estudo em relação ao total concedido	57,05%

Fonte: INSS

O resultado deste modelo será aplicado aos benefícios das categorias já citadas, atualmente concedidos e em manutenção, podendo implementar melhorias nos critérios de amostragem dos Órgãos de Auditoria do Governo Federal em relação à conformidade e exatidão dos benefícios concedidos ou em processo de concessão, segundo as regras descritas nos Anexos A, B e C.

Este modelo também deverá ser capaz de promover verificações no acervo mantido pela Previdência Social, buscando a conformidade dos benefícios pagos nestas categorias (Benefícios por incapacidade), podendo promover uma significativa redução dos valores pagos em função de incorreções nos critérios de concessão dos benefícios, embasando os processos de cessação e em caso de pagamentos já efetuados, possibilitar ao Ministério da Previdência e Assistência Social (MPAS) o desenvolvimento de ações de recuperação de valores pagos indevidamente, a partir da identificação dos benefícios e da execução de exames de conformidade detalhados, inclusive com avaliações presenciais dos segurados, além de outros processos de verificação definidos pelo Ministério da Previdência e Assistência Social (MPAS). O modelo resultante deste estudo poderá auxiliar na identificação e seleção de critérios dos benefícios, permitindo análises mais detalhadas em ocasiões definidas pelo INSS e pelos outros órgãos de controle do Governo Federal.

O presente modelo também poderá auxiliar na verificação da autenticidade e conformidade das concessões de Auxílios por Acidente de Trabalho, considerando as ocorrências históricas contidas nas bases de dados da Previdência Social como referencial. O modelo foi testado através da comparação dos resultados obtidos nos trabalhos anteriores de Auditoria sobre os benefícios integrantes da base histórica, a

partir dos sistemas de Tecnologia da Informação da DATAPREV – Empresa de Tecnologia e Informações da Previdência Social.

Em relação ao modelo de Previdência e Assistência Social vigente no Brasil Santos (2004, p. 29), afirma que:

Na maioria dos países, a forma de custeio da Seguridade Social tem características específicas, apesar de os benefícios oferecidos serem quase uniformes: aposentadoria única por velhice ou incapacidade permanente, pensão por morte e auxílio por incapacidade temporária (doença). Como exemplo, se compararmos cinco tipos de programas de seguridade social existentes no Brasil (velhice, incapacidade e morte, doença e maternidade, assistência médica e doenças do trabalho, desemprego e pensões familiares) com países da OCDE (Organização para a Cooperação e Desenvolvimento Econômico), teremos um percentual de quase 100% de coincidência, embora cada um desses países tenha suas particularidades.

Continuando, Santos (2004, p.30) ainda define como benefícios não-previsíveis os resultantes de acidentes ou doenças profissionais como:

No caso de incapacidade ou invalidez resultante de acidente ou doença ocasionada por fatores alheios às atividades laborativas, as pensões também são pagas, são os benefícios não-previsíveis, mesmo que a incapacidade seja permanente ou duradoura.

A relevância do estudo para a sociedade consistirá na ampliação da eficiência na análise dos processos de concessão e manutenção de benefícios, com possibilidades de resultados em economia de recursos públicos, gerados através de análise criteriosa do acervo de benefícios, que possibilitará melhorias na avaliação das concessões, manutenções ou dos cancelamentos dos benefícios, através da verificação mais abrangente da conformidade dos benefícios existentes na categoria de Benefícios por Incapacidade.

A ação contribuirá para a melhoria dos sistemas da Previdência Social através da incorporação da metodologia proposta aos sistemas existentes, bem como aos sistemas que integrarão o Novo Modelo de Gestão (NMG); contribuirá também para redução do *déficit* previdenciário e para a melhoria dos processos de

Auditoria da Previdência Social e da Dataprev, a partir de análises globais e definições de padrões e tendências sobre as informações previdenciárias; embasará as análises de conformidade e avaliações globais que serão mantidas permanentemente (*monitoramento contínuo* e *auditoria contínua*), em comparação com os procedimentos utilizados atualmente sobre amostragens e critérios legais.

O modelo gerado buscará também a redução dos valores de custeio relativos à Previdência Social, já que segurados que requerem aposentadorias antes da conclusão do prazo de contribuição e idade mínima (Benefícios não-programados), contribuem para o déficit da Previdência Social Brasileira, uma vez que o Sistema Previdenciário baseia-se em um fundo de repartição simples, ou seja, os contribuintes atuais financiam os beneficiários existentes no momento.

O investimento na melhoria da metodologia de acompanhamento e verificação de conformidade na concessão e manutenção dos benefícios das categorias de Auxílios por Acidente do Trabalho e Aposentadorias por Invalidez, proporcionará significativos avanços que poderão ser incorporados ao Novo Modelo de Gestão (NMG), permitindo também avanços nos processos de Auditoria e redução dos gastos públicos com o pagamento de benefícios pela Previdência Social Brasileira.

O resultado do estudo contribuirá para a melhoria da Gestão da Previdência Social, através da redução do seu déficit na área de Seguridade Social, proporcionando a melhoria na gestão dos pagamentos de benefícios por Incapacidade.

1.5 Hipóteses

De acordo com Kerlinger (1980, p. 38) apud Boaventura (2007), uma hipótese é “um enunciado conjectural das relações entre duas ou mais variáveis. São

sentenças declarativas e relacionam de alguma forma variáveis a variáveis”, continuando, Kerlinger (1980) apud Boaventura (2007) afirma que as hipóteses “São enunciados de relações, e, como os problemas, devem implicar na testagem das relações enunciadas”, afirmando também que “as hipóteses são instrumentos típicos das ciências físicas que vieram enriquecer as ciências sociais”.

Segundo Cooper e Schindler (2003, p. 57), uma hipótese é uma proposição formulada para testes empíricos, esta é feita como uma afirmação declaratória de natureza tentativa e conjectural.

Para Lakatos e Marconi (2008), as hipóteses são respostas prováveis ou provisórias para problemas que sejam cientificamente validados, sendo consideradas como colocações conjecturais da relação entre duas ou mais variáveis, devendo conduzir à implicações claras para o teste das relações existentes entre as variáveis e garantir que as variáveis sejam passíveis de mensuração.

Segundo Coen e Nagel (In: Selltitz et Alli, 1967:42-3) apud Lakatos e Marconi (2008), não é possível seguir adiante em uma pesquisa sem que as hipóteses sejam anunciadas, demonstrando uma solução provisória ao problema apresentado, sendo esta a função das hipóteses, “a busca de ordem entre os fatos”.

De acordo com Goode e Hatt (1968:74-5) apud Lakatos e Marconi (2008, p.144), as hipóteses formam os elos entre fatos e as teorias, a partir da reunião dos fatos, suas ordenações e relacionamentos e sob a orientação das hipóteses, válidas e sustentáveis, estas passam a integrar a teoria.

O papel das hipóteses, segundo Cooper e Schindler (2003, p. 58), é orientar a direção do estudo, identificar os fatos relevantes e não-relevantes, sugerir qual a forma de planejamento de pesquisa parece ser mais apropriada e fornecer uma

estrutura para organizá-las, além de sustentar e definir o que deve ser estudado, bem como serem passíveis de testes.

Abaixo estão relacionadas as hipóteses que motivam o estudo em questão:

- H_0 : Não há critérios comuns nos benefícios por incapacidade cessados que possibilitem a verificação da conformidade legal dos benefícios por incapacidade concedidos e mantidos.
- H_1 : A análise das características dos benefícios cessados, possibilitará a geração de Indicadores Antecedentes que serão aplicados na verificação de legitimidade dos benefícios mantidos pela Previdência Social.

1.6 Definição das variáveis

De acordo com Boaventura (2007), as variáveis constituem um dos principais fundamentos da ciência, afirma também que uma variável é algo que muda de valor, tornando-se diferente e que variáveis que possuam apenas um valor são designadas como constantes.

Segundo Kerlinger (1980) apud Boaventura (2007), as variáveis são propriedades que podem assumir valores diferentes, afirma também que as variáveis podem assumir muitos valores, porém estas devem assumir no mínimo dois valores.

Para Lakatos e Marconi (2008), a variável pode ser considerada uma classificação ou ainda uma medida, uma quantidade que varia, um conceito, um constructo que contém ou apresenta valores. Os valores que são acrescentados ao conceito podem ser quantidades, qualidades, características, magnitudes ou traços, que se alteram em cada caso particular. As autoras afirmam ainda que uma nova aplicação para a definição operacional é a transformação de constructos ou conceitos em variáveis, que caracterizam-se pela possibilidade de mensurá-las, em função de conterem valores ou classes de valores.

Tabela 3 - Definição inicial das variáveis utilizadas no estudo

Questão	Variável
1	Idade do Segurado na concessão do Benefício
2	Idade do Segurado em Janeiro/ 2008
3	Sexo do Segurado
4	Tempo de Contribuição
5	Espécie do Benefício
6	Forma de Filiação ao Regime Geral de Previdência Social (RGPS)
7	Atividade desempenhada pelo segurado (CBO)
8	Quantidade de dependentes para Imposto de Renda (IR)
9	Renda Mensal Inicial prevista em Salários-Mínimos
10	Tipo de acidente
11	Ramo de atividade
12	Quantidade de revisões da Renda Mensal Inicial
13	Se tem Perícia Médica realizada
14	Unidade da Federação onde o benefício foi concedido
15	Unidade da Federação onde o segurado trabalhou
16	Unidade da Federação onde o benefício é recebido
17	Tipo de Sistema de Concessão
18	Uso de informações do CNIS
19	Tipo de despacho do benefício
20	Tempo de Recebimento do Benefício

Fonte: Elaboração própria a partir de entrevistas com especialistas da Previdência Social

A tabela 3 apresenta as variáveis que delimitaram o estudo, sendo possível a integração de novas variáveis a partir de consultas aos funcionários especialistas da Previdência Social e da DATAPREV.

Após 5 (cinco) reuniões, com os especialistas da Área de Gestão e Manutenção de Benefícios da Dataprev, foi obtido o conjunto das variáveis que na opinião dos especialistas, deveria ser analisado com o propósito de identificar o perfil dos beneficiários que mantiveram em algum momento um vínculo com a Previdência Social através de um Benefício por Incapacidade, mesmo que este benefício esteja na condição de cessado no mês de janeiro do ano de 2008, data limite que este estudo considera como referência para a avaliação dos benefícios. Como resultado das entrevistas com o corpo funcional e especialistas na área de Previdência Social, foram identificadas as seguintes variáveis:

ID_SEG_DIB – Idade do segurado no momento da requisição do benefício por incapacidade e do respectivo início do benefício.

ID_SEG_2008 – Idade do segurado em janeiro de 2008 (data que limitará a avaliação promovida pelo presente estudo).

CS_SEXO – Código atribuído pela Previdência Social Brasileira para indicar o sexo do segurado. A informação relativa ao sexo dos beneficiários possui alta relevância nos processos de transformação de benefícios relativos aos Auxílios por Acidente em Aposentadorias por Incapacidade, em tal situação os prazos considerados para que o segurado tenha direito ao valor de benefício integral varia em função do sexo do segurado (30 ou 35 anos) excetuando-se categorias especiais, nas quais há períodos de contribuição de menor duração.

TP_CONTRIB – Tempo de contribuição do segurado junto à Previdência Social desde a sua filiação ao RGPS até o momento da solicitação do benefício por incapacidade, considerando as interrupções nas contribuições desde que não excedam aos períodos definidos para que o cidadão permaneça como segurado, mesmo que haja interrupção nas contribuições ou na configuração de vínculos empregatícios.

CS_ESPECIE – Classificador numérico de utilização interna pela Previdência Social para definir os tipos de benefícios abrangidos pela Previdência Social aos seus segurados.

CS_FORMA_FILIACAO – Classificador numérico utilizado pela Previdência Social para categorizar os tipos de vínculos da filiação do segurado ao RGPS (Regime Geral de Previdência Social).

CS_CBO – Classificador utilizado pela Previdência Social para identificar a atividade profissional desempenhada pelo segurado (Classificação Brasileira de Ocupações), segundo as definições do Ministério do Trabalho e Emprego, disponíveis em <http://www.mtecbo.gov.br/busca.asp> (acessado em 04/07/2008).

QT_DEP_IR – Quantidade de dependentes para Imposto de Renda informada pelo Segurado à Previdência Social e vigente em seu vínculo com o RGPS, considerando as regras de dependência aceitas pela legislação fiscal vigente no Brasil.

QT_SM_RMI – Quantidade de Salários Mínimos relativa ao valor do benefícios recebido pela Previdência Social em decorrência de uma incapacidade para exercer atividade profissional (Renda Mensal Inicial).

CS_ACIDENTE – Classificador utilizado pela Previdência Social para retratar o tipo de acidente que motivou a concessão do benefício por Incapacidade ora prestado.

CS_RAMO_ATIVIDADE – Classificador utilizado pela Previdência Social para retratar o setor empresarial no qual o segurado exerce a atividade profissional.

QT_REVISAO_RMI – Quantidade de revisões solicitadas pelo segurado ou promovidas automaticamente pela Previdência Social no valor definido como a renda que o segurado recebe mensalmente (RMI – Renda Mensal Inicial), apurada nos processos de concessão, decisões judiciais ou ainda revisões administrativas promovidas pela Previdência Social, visando a adequação das regras de negócio às alterações no sistema de Previdência Social.

IN_PERICIA_MEDICA – Indicador da realização de uma Perícia realizada através de um médico ou junta médica da Previdência Social ou conveniada a esta, para verificar ocorrência do acidente ao segurado, a sua capacidade laboral ou ainda o tempo necessário para a sua recuperação, quando uma nova perícia deverá ser realizada, avaliando a sua recuperação ou um novo período de recuperação ou a concessão de sua aposentadoria em caráter definitivo devido à impossibilidade do segurado continuar desenvolvendo atividade profissional.

CS_ULTIMO_LT – Classificador utilizado pela Previdência Social para retratar o último local (UF - Unidade da Federação) em que o segurado desenvolveu suas

atividades profissionais remuneradas. Esta informação permite à Previdência Social avaliar a movimentação dos trabalhadores e das atividades e categorias funcionais que são desenvolvidas nas Regiões Geográficas Brasileiras, além da capacidade de absorção de mão-de-obra ao longo do Território Nacional nas diversas áreas de atividades trabalhistas.

CS_SISTEMA_CONCES - Classificador utilizado pela Previdência Social para retratar o sistema de concessão utilizado para conceder o benefício por incapacidade ao segurado, este classificador corresponde às diversas formas de solicitação do benefício (internet, Agência da Previdência Social, dentre outros) ao longo do tempo. Através deste indicador, a Previdência Social avalia desempenhos e particularidades existentes nos seus sistemas (tecnologia e gestão) ao longo da trajetória do segurado, considerando informações e regras vigentes ao longo do tempo e suportadas pelo RGPS (Regime Geral de Previdência Social).

CS_BATIMENTO_CNIS – Classificador que retrata o resultado da verificação da situação do segurado em relação aos vínculos e remunerações mantidas pelo segurado, decorrentes de sua atividade profissional remunerada. As informações sobre os vínculos e remunerações dos segurados são mantidas no Cadastro Nacional de Informações Sociais (CNIS), esta base de informações é mantida e atualizada a partir de informações de vínculos e remunerações enviadas pela Caixa Econômica Federal (CEF) e por informações demográficas repassadas pelo Ministério do Trabalho e Emprego. As informações são gerenciadas pela DATAPREV (Empresa de Tecnologia e Informações da Previdência Social), que através de definição legal, somente poderá conceder benefícios após examinar a situação dos segurados junto ao CNIS, evitando sempre que possível o “ônus da

prova” (o segurado só apresentará documentação quando necessário ao processo de concessão de benefícios).

CS_MOT_DESP – Classificador utilizado pela Previdência Social para identificar os depachos de natureza administrativa (sob a análise do INSS) ou judicial (resultante de ações judiciais impetradas por segurados em circunstâncias que necessitam de análise mediadora entre o plano de benefícios e o entendimento dos segurados solicitantes dos benefícios ou as alterações nos mesmos).

QT_ANOS_RECEBE – Quantidade de anos que o beneficiário recebeu ou recebe o benefício por incapacidade garantido pelo Regime Geral de Previdência Social (RGPS).

2 REFERENCIAL TEÓRICO

2.1 Análise projetiva

O fascínio da humanidade com o futuro pode ser remontado aos tempos pré-históricos. Há evidências concretas mostrando que as pessoas desejam prever o futuro na intenção de obter vantagem, fama e poder. Há base científica na previsão (Análise Projetiva), mas também há sérias limitações na capacidade de predição de fatos e situações futuras (MAKRIDAKIS, 1990).

A preocupação permanente dos seres humanos com o futuro e suas implicações sobre as suas decisões atuais, levaram durante muito tempo as sociedades a acreditar na influência puramente da natureza sobre o desenrolar das ações ao longo do tempo. As sociedades conseguiram realizar as projeções para o futuro, baseadas em situações do passado e do presente, em médio e curto prazo, de forma intuitiva. Diante da incerteza de concretização do futuro e da agilidade requerida das decisões, dos riscos aos quais as decisões estão submetidas, da assimetria de informações e dos impactos resultantes sobre as decisões do presente, a análise projetiva assume relevante importância nas sociedades modernas e incorpora fatores métricos, demográficos, econômicos e sociais na formulação de seus modelos projetivos (BUARQUE, 2003).

As previsões têm o propósito de tornar um pensamento público, sendo fundamentadas por três razões: satisfazer a curiosidade, embasar os tomadores de decisão e gerar consenso entre as pessoas. A previsão como uma forma de consenso está ligada às formas de gerenciamento das organizações e suas

movimentações estratégicas, através de cenários futuros e previsões de longo prazo (MAKRIDAKIS, 1990, p. 55).

Segundo Schwartz (2000), os cenários concebidos através dos enfoques probabilísticos podem ser classificados em Cenário Projetivo e Cenário Prospectivo. O Cenário Projetivo tem uma única visão de futuro, destina-se predominantemente aos ambientes de baixa ou pouca turbulência e são aconselháveis para projeções de curto prazo, realizando projeções baseadas no comportamento do passado e do presente, considerando que não ocorrerão fatos que alterarão as tendências identificadas. Em relação aos Cenários Prospectivos, possibilitam diversas visões do futuro, projetam tendências, procuram captar rupturas e descontinuidades, adaptando-se melhor aos ambientes turbulentos e ao de longo prazo.

A descoberta de conhecimento em bases de informações é um campo de pesquisa que tem crescido rapidamente, tem sido direcionado em benefício de diversas áreas de aplicações (sociais, econômicas, humanas e outras). Na atualidade, os grandes volumes de dados armazenados conduzem a uma poderosa tecnologia de coleta. As informações armazenadas, em sua grande maioria, possuem valiosas relações sobre tendências e padrões que podem ser utilizados para melhorar as decisões de negócios (REZENDE, 2003).

Para Moritz (2004), o ser humano é essencialmente limitado e não é capaz de adivinhar o futuro, porém, é capaz de explorar configurações futuras de suas variáveis ambientais mais relevantes. Através da combinação entre ciência e arte e do uso de instrumental matemático, torna-se possível esclarecer as relações quantitativas e qualitativas entre as variáveis, embora o papel preponderante no processo cabe ao perito ou especialista, que deve ser capaz de utilizar

conhecimentos teóricos e práticos, além da sua sensibilidade, para elaborar visões coerentes do futuro.

De acordo com Simon (1957), um dos pioneiros nos estudos sobre processo decisório, escreveu sobre a psicologia nas decisões administrativas, abordando a racionalidade na tomada de decisões, deixando claras as imperfeições do conhecimento que afetam o raciocínio humano. Segundo Simon (1957 p. 84)

“... a racionalidade requer um conhecimento completo e inalcançável das conseqüências exatas de cada escolha. Na verdade o ser humano possui apenas um conhecimento fragmentado das condições que cercam sua ação, e ligeira percepção das regularidades dos fenômenos e das leis que lhe permitiriam gerar futuras conseqüências com base no conhecimento das circunstâncias atuais.”

Para Simon (1957) o ser humano age na maioria das vezes sob condições de racionalidade limitada, uma vez que os humanos têm dificuldades de considerar e avaliar todas as informações e variáveis envolvidas em um processo de decisão, sendo feita a escolha de apenas algumas variáveis para tomar a decisão necessária.

Para WANG (2008), os dados são os ativos mais importantes para a realização de análises detalhadas, essenciais para a tomada de decisões corretas e competentes em um mundo de constantes mudanças no ambiente empresarial. Embora as empresas possuam mecanismos de análise sobre os seus dados, o grande crescimento das informações armazenadas conduzem a uma análise com um grande atraso sobre as informações em poder das corporações. As corporações que detêm a posse dos dados, possuem especialistas que dominam os processos que geram os dados e o seu significado para as atividades das corporações.

Segundo Gardner (1998 apud REZENDE, 2003) a realização de um *Data Warehouse* é um dos primeiros passos para tornar factível a análise de grandes volumes de dados no apoio ao processo decisório. O objetivo do *Data Warehouse* é criar um repositório de dados limpos, agregados e consolidados que possam ser

analisados por ferramentas OLAP (*On-Line Analytical Processing*). Estas ferramentas apresentam facilidades para a realização de consultas complexas em bases de dados multidimensionais. São destinadas aos usuários para que estes realizem as consultas, visando atender às suas hipóteses pendentes de comprovações, através dos resultados das consultas realizadas à base de conhecimento multidimensional.

Diante da deficiência para analisar e compreender um grande volume de dados, diversos estudos têm sido direcionados ao desenvolvimento de tecnologias de extração automática de conhecimento de bases de dados. Esse campo da pesquisa é chamado de Extração de conhecimento de bases de dados ou *Knowledge Discovery in Database* (KDD) ou Mineração de dados ou *DataMining* (DM) que são considerados processos distintos (FAYYAD, PIATETSKY-SHAPIRO & SMYTH, 1996 apud REZENDE, 2003), incorporando técnicas das áreas de Base de dados, Inteligência Artificial e Estatística, sendo o conhecimento alvo das descobertas sobre as bases de dados, compreensíveis aos seres humanos e úteis no embasamento dos processos de tomada de decisão.

Para Fayyad, Piatetsky-Shapiro & Smyth (1996a) apud Rezende (2003), a Extração de conhecimento de base de dados é “o processo de identificação de padrões válidos, novos, potencialmente úteis e compreensíveis embutidos nos dados”.

De acordo com Simon (1960), o papel do gestor é tomar decisões. O exercício deste papel não se resume apenas ao ato final da escolha entre as alternativas disponíveis, mas todo o processo que inclui: descobrir as ocasiões em que deve ser tomada a decisão, identificar os possíveis cursos de ação e decidir-se entre um deles.

Simon (1960) ainda aborda instrumentos matemáticos que podem ser utilizados em decisões programadas, com o objetivo de tornar quantitativas situações em que não existem mensurações, destacando assim, a importância da tomada de decisão racional no ambiente organizacional.

Os sistemas de mineração de dados são desenvolvidos para os mais variados domínios. Essas tarefas podem extrair diferentes tipos de conhecimento, sendo necessário decidir já no início do processo de *DataMining*, qual o tipo de conhecimento que o algoritmo deve extrair. As atividades de *predição* envolvem o uso dos atributos de um conjunto de dados para prever o valor futuro da variável meta, buscando apoio para a tomada de decisão, enquanto as atividades de *descrição* procuram padrões interpretáveis pelos humanos que descrevam os dados antes de realizar a previsão.

Uma regra de *associação* caracteriza o quanto a presença de um conjunto de atributos nos registros de uma base de dados implica a presença de algum outro conjunto distinto de atributos nos mesmos registros (AGRAWAL e SRIKANT, 1994 apud REZENDE, 2003). Para Fayyad, Piatetsky-Shapiro e Smyth, 1996b apud Rezende (2003), a sumarização envolve métodos para encontrar uma descrição para um subconjunto de dados. Um exemplo de sumarização é a definição da média e do desvio-padrão de todos os campos.

De acordo com MAKRIDAKIS (1990), futuro pode ser previsto através da observação das regularidades percebidas no passado e no presente (padrões – *patterns*), ou das relações de causalidade. Uma premissa para a previsão, seja estatística ou de julgamento, é a de que o padrão ou o relacionamento exista e seja percebido no evento alvo do estudo, sendo este corretamente identificado, poderá ser projetado, permitindo a previsão.

Para MAKRIDAKIS (1990), as formas de previsão (Análise Projetiva) seguem critérios para ter o seu resultado aceito cientificamente, tais como:

- Os dados são identificados e coletados;
- Os modelos estatísticos são selecionados;
- Os padrões e relacionamentos dos dados envolvidos são identificados e criteriosamente mensurados;
- As previsões são realizadas através da projeção dos padrões e relacionamentos identificados e mensurados;
- As incertezas das previsões podem ser estimadas e utilizadas para reduzir as conseqüências negativas dos eventos inesperados.

A precisão das previsões realizadas nas áreas das ciências naturais e físicas, ocorre em função dos fenômenos práticos inerentes a estas ciências, enquanto que a mesma precisão não deve ser atribuída às outras áreas da ciência, tais como a precisão nas previsões sobre fenômenos das áreas sociais, tais como economia e negócios, nas quais a previsão torna-se muito mais sujeita a fatores externos e influência de mudanças nos padrões e relacionamentos (MADRIDAKIS,1990).

Dentre os diversos fatores que poderão alterar a precisão dos processos de análise projetiva, estão: o horizonte de tempo, as mudanças tecnológicas, a disseminação da informação e as mudanças sociais e econômicas que influenciam hábitos e padrões comportamentais.

Para Rezende (2003), a modelagem de dependência consiste em encontrar um modelo que descreva as dependências mais significativas entre as variáveis. Há dois níveis de modelos de dependência: o nível *estrutural* que especifica o modelo (normalmente através de forma gráfica), no qual as variáveis são dependentes

localmente, e o nível quantitativo do modelo que especifica o peso das dependências usando uma escala numérica.

A análise projetiva, parte do pressuposto de que o futuro é visto como uma continuação das tendências passadas e analisado como único e certo.

Os cenários preditivos são obtidos pela extrapolação dos fatos passados e consideram que as forças que moldaram o passado e construíram o presente continuarão a atuar e modelar o futuro.

Para Börjeson et al. (2005), o cenário projetivo mostra o que irá ocorrer, o cenário prospectivo expõe o que pode ocorrer, enquanto o cenário normativo indica como atingir um objetivo específico.

A construção do futuro esta centrada nas forças que moldam os acontecimentos e modelarão o futuro, a partir das quais podem ser traçados dois tipos de cenários preditivos: Previsão (*Forecasts*) e o que fazer se (*What-if*) um determinado fato ou uma consequência influenciar os acontecimentos futuros, mesmo não prevista no momento do planejamento.

Deve ser feita uma distinção entre a abordagem projetiva e a prospectiva. A primeira (projetiva) considera que as forças que modelaram o passado continuarão a modelar o futuro, situação ilustrada na figura 2. Este, portanto, poderá ser previsto, e será único.

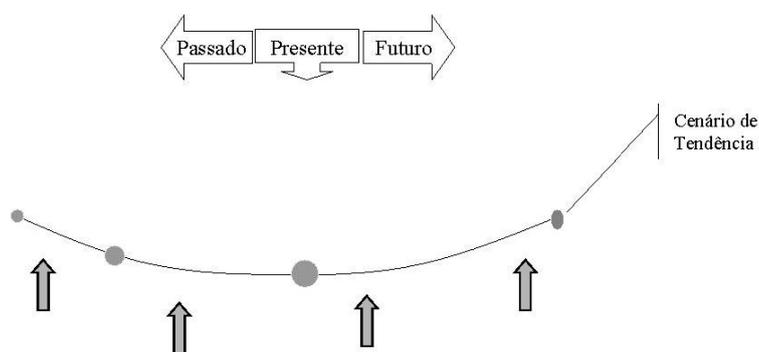


Figura 2 - Demonstração dos estágios da análise projetiva

No caso de uma análise projetiva, a ruptura de tendências conduzirá a outro cenário, diverso do primeiro, porém igualmente único e definido.

As técnicas de prospecção de futuro, segundo Makridakis e al. (1983), são divididas em métodos quantitativos e qualitativos.

Os métodos quantitativos são aplicados quando:

1. Existe informação suficiente sobre o passado;
2. A informação pode ser quantificada em forma de dados numéricos;
3. Pode-se assumir que alguns aspectos dos padrões do passado continuarão no futuro.

Esta última condição, chamada de “pressuposto da continuidade”, aparece em todos os modelos quantitativos e em alguns modelos qualitativos (MAKRIDAKIS e al., 1983).

As áreas de estatística e mineração de dados estão fortemente ligadas, já que ambas têm como objetivo encontrar padrões e regularidades nos dados, além do fato de que a maioria dos algoritmos utilizados está voltada para a produção de conjuntos de declarações sobre dependências locais entre variáveis de interesse do que propriamente para a construção de modelos globais que incluam todas as variáveis de interesse do problema (GLYMOUR, MADIGAN, PREGIBON e SMYTH, 1997 apud REZENDE, 2003).

2.2 Indicadores Antecedentes nas diversas áreas

Com o propósito de embasar a utilização da regressão logística na condução do estudo, serão apresentados vários estudos realizados no âmbito acadêmico e empresarial que utilizaram a técnica multivariada de regressão logística nas diversas áreas do conhecimento. Os estudos utilizaram-se da análise projetiva para construir

indicadores antecedentes, com o propósito de elaborar previsões e descrever as constatações realizadas através da análise de ocorrências históricas e atuais.

Para GUIMARÃES e NETO (2002), a inadimplência é um dos maiores problemas, enfrentado pelas administradoras de cartão de crédito. Em seu estudo foi avaliado o conceito de risco, que é essencialmente a probabilidade de não recebimento dos créditos por parte das administradoras de créditos e cartões. Caouette et al. (2000) e Silva (1988) apud Guimarães e Neto (2002), referem-se às técnicas estatísticas multivariadas como ferramentas poderosas na administração do risco envolvido na concessão de crédito pessoal.

O trabalho de Guimarães e Neto (2002) apresenta a construção e avaliação de regras de reconhecimento de padrões baseadas em duas técnicas multivariadas: a Função Discriminante Linear de Fisher e a Regressão Logística para classificação de clientes de cartões de crédito em um de dois grupos (adimplentes ou Inadimplentes).

Através da análise das concessões de crédito, via cartão de crédito, por ser este um instrumento de crédito automático, tornou-se obrigatória e permanente a busca por técnicas que permitam o gerenciamento de um grande número de empréstimos aos diversos perfís de consumidores, visando a obtenção do maior retorno possível para as instituições de crédito a partir dos créditos disponibilizados aos consumidores (GUIMARÃES E NETO (2002).

Capelleto, Martins e Corrar (2008) desenvolveram um modelo utilizando a regressão logística para realizar a previsão do risco sistêmico do setor bancário a partir de variáveis econômicas e contábeis, utilizando um estudo indutivo, coletando as informações financeiras de instituições localizadas em 30 (trinta) países onde ocorreram crises financeiras a partir de 1990. Foram selecionados

indicadores sobre a solidez econômica e financeira das instituições bancárias, bem como de solidez do ambiente econômico. Os testes estatísticos realizados com a regressão logística comprovaram a existência de indicadores contábeis e de riscos capazes de discriminar os sistemas bancários como suscetíveis ou não-suscetíveis às crises. As equações apresentaram percentuais de acerto nas classificações superiores a 90%.

A partir do modelo, tornou-se possível a correta separação dos grupos. As classificações dos países são estabelecidas em percentuais e expressam a probabilidade de pertencer a determinado grupo. O ordenamento dos países pelo grau de risco sistêmico fornece parâmetros de comparação e propicia a tomada de decisão calibrada à exigência de cada situação. Por meio dele, é possível saber qual país contém maior risco sistêmico no setor bancário.

As variáveis contábeis e econômicas mais associadas à ocorrência de crises estão relacionadas com a qualidade dos créditos, o volume de resultados e o nível de taxa de juros. Todos os indicadores formados com essas variáveis foram relevantes no processo de classificação, destacando-se os referentes à volatilidade da inadimplência, da rentabilidade e da taxa de juros, e à média da rentabilidade e do risco de crédito.

Em estudo realizado por Carvalho e Cortes (2000) foram analisados através da aplicação de Regressão Logística, a mobilidade ocupacional entre as raças, com enfoque nas origens, destinos e riscos dos Afro-Brasileiros, através da estimativa do risco inerente às atividades ocupacionais por grupo racial em diferentes níveis de ocupações no mercado de trabalho, através do cálculo de probabilidade da saída das posições ocupadas pelos grupos étnicos. A pesquisa baseou-se nas informações da Pesquisa Mensal de Emprego (PME) nas seis principais regiões

metropolitanas brasileiras, realizado pelo IBGE no ano de 1996. Foi avaliada a importância da raça na determinação de movimentos ocupacionais, considerando as características como capital humano (escolaridade do indivíduo, do pai, da mãe, qualidade da educação e experiência aproximada por idade), a natureza da inserção produtiva (região, setor de atividade e acesso a novas tecnologias) e outros atributos pessoais (sexo, posição na família e religião). Os autores concluíram que as transições das posições mais altas para as mais inferiores são positivamente correlacionadas com os grupos não Afro-Brasileiros (Branços e Amarelos) e que os não Afro-Brasileiros ocupantes de posições altamente qualificadas apresentam 93,79% de chances de precarizar suas ocupações, quando comparados com Afro-Brasileiros.

Segundo Almeida e Siqueira (1997), a avaliação de risco de insolvência de bancos brasileiros é uma questão particularmente importante no contexto econômico. A avaliação de riscos de insolvência tem sido extensamente efetuada a partir da década de 60, essencialmente voltada para o segmento empresarial não-financeiro (COHEN, 1966; BEAVER, 1966; ALTMAN, 1968; ABATE, 1969 apud Almeida e Siqueira,1997), fundamentados no uso de regressão e análise discriminante e classificatória. Há outros estudos voltados para a utilização de redes neurais na avaliação de riscos de insolvência, ainda no segmento não-financeiro das economias (ALMEIDA,1993; JOHNSEN, 1994; WILSON, 1994; PODDIG,1995 apud ALMEIDA e SIQUEIRA,1997).

A maioria dos modelos de avaliação de riscos de insolvência é construída usando-se amostras pareadas: uma parte da amostra contém informações sobre empresas insolventes; a outra parte contém informações sobre empresas não insolventes. Variáveis são selecionadas segundo seu interesse potencial na

avaliação de riscos de insolvência. Um método estatístico é então usado para desenvolver um modelo discriminante (i.e., uma combinação de variáveis que melhor distinga os dois tipos de empresa). Finalmente o sucesso da discriminação é avaliado através de uma amostra de controle (i.e., uma amostra diferente da usada para obter o modelo).

Em geral a amostra de controle contém empresas solventes e empresas comprovadamente insolventes, isto é, empresas que estão em processo de concordata, em liquidação, ou no caso de bancos e instituições financeiras, sob intervenção do Banco Central.

Desde OHLSON (1980) a análise LOGIT é usada freqüentemente para a avaliação de riscos de inadimplência, baseada em características financeiras (índices) das empresas. O modelo LOGIT cria para cada empresa um score Z:

$$Z = a + \beta X_i$$

onde X_i é o valor da i -ésima variável (i.e. características financeiras)

$$P = 1 / (1 + e^{-(a + \beta X_i)})$$

Como P sempre se situa entre 0 e 1, o score Z é freqüentemente interpretado como a probabilidade de inadimplência.

No estudo promovido por Almeida e Siqueira (1997), foram considerados para a análise de risco de bancos, 9 indicadores financeiros, descritos nos Manuais de Análise da Austin Asis, empresa brasileira de análise de risco bancário. Além de tais indicadores, foi considerada a variável de indicação de porte $\log(\text{Ativo})$. Para a criação de modelos de análise de risco de insolvência, dois grupos são necessários. Um primeiro grupo contendo instituições insolventes e um segundo grupo contendo instituições solventes. As instituições consideradas insolventes neste estudo são aquelas que sofreram processos de liquidação ou intervenção do Banco Central do

Brasil. Para o grupo de solventes, foi escolhido um grupo de bancos considerados tradicionais (27 bancos).

O uso de regressão logística também foi comprovado por FRÁGUAS JUNIOR et ALVES (2002), em estudo sobre os transtornos depressivos e o seu grande prejuízo na esfera pessoal e familiar, acometendo principalmente adultos jovens – entre 20 e 30 anos de idade, atingindo duas vezes mais mulheres que homens. O *National Comorbidity Survey* relatou uma prevalência de depressão maior ao longo da vida de 17,1%.

O objetivo do estudo de FRÁGUAS JUNIOR e ALVES (2002) foi investigar os sintomas da depressão associada a condições médicas no Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HCFMUSP) e identificar aqueles que diferenciam pacientes com e sem depressão. Foram estudados, de modo consecutivo, 293 casos atendidos, no ano de 1998, pelo Grupo de Interconsultas do Instituto de Psiquiátrica do HCFMUSP.

Para GIMENO e SOUZA (1995), em epidemiologia, a regressão logística tem como objetivo descrever a relação entre um resultado (variável dependente ou resposta) e um conjunto simultâneo de variáveis explicativas (preditoras ou independentes), mediante um modelo que tenha bom ajuste, que seja biologicamente plausível e obedeça ao princípio da parcimônia. Na análise estratificada tem-se o mesmo propósito, mas as relações são efetuadas uma a uma, isto é, somente é possível obter a estimativa do risco para um único fator de cada vez, controlando-se o conjunto das demais variáveis.

Ainda segundo GIMENO e SOUZA (1995), podem ser citadas como vantagens da análise estratificada sua relativa simplicidade de execução, a facilidade de entendimento e maior proximidade que propicia entre o pesquisador e

os dados. Entretanto, ao se aplicar essa técnica, o grau de dificuldade aumenta na medida em que aumenta o número de variáveis que tiverem que ser consideradas como controle; os testes de homogeneidade entre os estratos, para se verificar a existência de interação entre as variáveis, são feitos em etapa à parte; variáveis quantitativas não podem ser usadas em sua escala original e o nível global de significância é difícil de ser controlado.

De acordo com GIMENO e SOUZA (1995), a análise logística controla grande número de variáveis simultaneamente, permitindo que os dados sejam utilizados mais eficientemente; o teste de homogeneidade pode ser feito em conjunto, bastando introduzir no modelo o termo produto entre os fatores. Uma possível desvantagem é a eventual barreira que introduz entre o pesquisador e os dados; é praticamente obrigatório o uso de pacotes estatísticos e microcomputador.

O trabalho de GIMENO e SOUZA (1995) exemplifica a utilização das análises estratificada e logística na análise de dados de estudos tipo caso-controle, através de estudo caso-controle sobre câncer de esôfago. Oitenta e cinco casos e 292 controles foram classificados segundo sexo, idade e os hábitos de beber e de fumar. As estimativas por ponto dos *odds ratios* foram semelhantes, sendo as duas técnicas consideradas complementares.

O hábito de beber foi considerado fator de risco de principal interesse. Foi verificada a existência de associação entre o câncer de esôfago e cada uma das variáveis, em uma primeira etapa (análise bruta), mediante a obtenção das estimativas dos *odds ratios* (OR), por ponto e por intervalo, além do valor da estatística qui-quadrado de Mantel-Haenszel. Nas etapas seguintes foram selecionadas as variáveis cujo valor do nível descritivo de significância do teste fosse menor do que 0,205 (GIMENO e SOUZA, 1995).

O trabalho desenvolvido por SANTOS₁ et. al. (2008), teve como objetivo elaborar um sistema para predição da soroprevalência da Hepatite A. Para isto, foram considerados os modelos de regressão de logística e redes neurais.

O principal objetivo deste trabalho foi a criação de um sistema de predição da soroprevalência da hepatite A, visando o apoio ao diagnóstico. O sistema a partir de uma base de informações relevantes para o problema em estudo, contribuiu na identificação de indivíduos com alto risco de contrair a doença, atenuando o risco de disseminação da doença para a população, como também na identificação daqueles indivíduos que mereceram posterior investigação em Unidades de Saúde.

A amostragem dos indivíduos foi estratificada por grupo etário, tendo sido selecionados 3.079 indivíduos com idade entre 1 e 83 anos.

Os indivíduos foram classificados como soropositivo ou soronegativo e as variáveis relacionadas ao indivíduo, seu ambiente domiciliar e peridomiciliar, foram coletadas por meio de entrevistas domiciliares, através de um formulário pré-codificado. Foram observadas 66 variáveis, entre as quais: idade, sexo, renda mensal e escolaridade da dona de casa, a não utilização de filtro de água, densidade de moradores por cômodos, tempo de moradia na residência, condições sanitárias e outras. Foi criado um modelo logístico com as 66 variáveis explicativas observadas nos 3.079 indivíduos residentes nos domicílios amostrados. Em seguida, através deste modelo, foram selecionadas as variáveis independentes que foram estatisticamente significantes.

Através do modelo de regressão logística, foi possível identificar, por exemplo, as variáveis relacionadas ao ambiente domiciliar e peridomiciliar que estão associadas a uma maior ou menor probabilidade de se adquirir a doença.

As sete variáveis independentes resultantes foram: idade, proximidade do domicílio à vala negra, densidade de moradores/cômodos, número de pontos de água no domicílio, a não utilização de filtro de água, número de anos de estudo e renda média mensal da dona de casa.

Segundo ONUSIC, KASSAI e VIANA (2004), a análise de indicadores extraídos das demonstrações financeiras colabora para a tomada de decisão sobre investimento e empréstimo, pois auxilia na avaliação do grau de solvência de uma empresa. O objetivo da pesquisa foi desenvolver um modelo de classificação de risco para avaliar o risco de crédito de empresas no mercado brasileiro, de capital aberto e classificadas como solventes ou insolventes no período entre 1994 e 2004, através da aplicação da análise multivariada, utilizando a Regressão Logística. As variáveis independentes são índices financeiros calculados a partir das demonstrações contábeis e utilizados para representar a situação econômico-financeira das empresas. Os resultados do estudo indicam que o modelo de classificação de risco desenvolvido prevê eventos de *default* com um ano de antecedência com bom nível de acurácia. Os resultados, também, indicam que as demonstrações contábeis contêm informações que possibilitam a classificação das empresas como prováveis solventes ou prováveis insolventes.

Segundo Bessis (1998, p.81) apud ONUSIC et al. (2004), o risco de crédito pode ser definido pelas perdas geradas por um evento de *default* do tomador ou pela deterioração da sua qualidade de crédito. A mensuração de risco de crédito é o processo de quantificar a possibilidade de a instituição financeira incorrer em perdas, caso os fluxos de caixa esperados com as operações de crédito não se confirmem.

O risco de *default* constitui a principal variável desse processo, podendo ser definido como a incerteza em relação à capacidade de o devedor honrar os seus compromissos assumidos (BRITO e ASSAF NETO, 2008).

De acordo com BRITO e ASSAF NETO (2008), os modelos que buscam prever a insolvência de empresas têm sido estudados há várias décadas nos meios acadêmicos, e há na literatura um conjunto de pesquisas que se tornaram referências, os trabalhos de Beaver (1966), Altman(1968) e Ohlson (1980) em âmbito internacional, e os de Kanitz(1976), Matias (1978) e Altman et al. (1979) no Brasil.

A amostra final utilizada no desenvolvimento do modelo compreende 60 empresas, sendo 30 insolventes que se tornaram concordatárias ou falidas entre 1994 e 2004 e 30 solventes que foram emparelhadas com as primeiras. Esse método de emparelhamento da amostra foi baseado nos trabalhos de Beaver (1966), Altman (1968) e Sanvicente e Minardi (1998), podendo ser considerado como uma amostragem por julgamento.

As variáveis explicativas do modelo de classificação de risco de crédito compreendem índices utilizados para caracterizar a situação econômico-financeira das empresas, calculados a partir das suas demonstrações contábeis publicadas. O modelo proposto nesta pesquisa está fundamentado na relação entre a situação econômico-financeira da empresa e a ocorrência do evento de *default*. Dessa forma, quanto pior for a situação econômico-financeira de uma empresa, maior será a probabilidade de ela incorrer em *default*.

Segundo Hosmer e Lemeshow (1989), a validação do modelo é especialmente importante quando ele é usado com a finalidade de previsão de resultados.

O trabalho de Pereira e Ness-Júnior (2008) utilizou a regressão logística para estabelecer um modelo estatístico e evidenciar a probabilidade de ocorrência de falências ou concordatas em empresas de internet atuantes no mercado dos Estados Unidos. O *e-score* foi definido através das relações entre as variáveis obtidas a partir das informações publicadas sobre o desempenho das empresas do setor. Através de uma amostra por conveniência sobre as empresas constantes no *American Stock Exchange* no ano de 2002, foram consideradas 36 empresas não-falidas e 25 empresas falidas ou concordatárias.

Foram considerados os índices constantes nos referenciais bibliográficos sobre o tema e incluídos índices que somente devem ser considerados em empresas que atuam exclusivamente no âmbito de internet (atividade única). Foi utilizado o teste de *Levene* das variáveis a um nível de significância de 5%. Após este teste foi realizado o teste de *Kolmogorov-Smirnov* para verificar a normalidade da distribuição dos dados, possibilitando a adoção da regressão logística, como um modelo não-paramétrico, já que somente 4 variáveis apresentaram um padrão de normalidade, ao nível de significância de 5%. Foi utilizado o software estatístico versão 10.0 na aplicação do modelo *logit*, com o propósito de analisar as variáveis explicativas que melhor separam as empresas nos grupos de falidas ou não-falidas. Foram identificadas 3 variáveis explicativas que possibilitam classificar com precisão de 97,4% as empresas entre os dois grupos (falidas e não-falidas) na amostra de treinamento (60% do total). Já na amostra de teste (40% do total) o nível de precisão foi de 91,3%, sendo esta definida como a precisão final do modelo.

O trabalho de LAVINAS e NICOLL (2006) avaliou através da identificação de variáveis explicativas, utilizando a técnica multivariada de Regressão Logística, os

fatores que em maior ou menor grau evidenciaram a vulnerabilidade social, buscando relações entre os níveis de pobreza e desigualdade salarial entre homens e mulheres no mercado de trabalho. Com base na PNAD (2003) realizada pelo IBGE, os autores analisaram as diferentes participações no mercado de trabalho em função do gênero dos trabalhadores, os diferentes rendimentos e as diferentes oportunidades apresentadas pela pesquisa em relação aos homens e mulheres investigados no nível nacional. O estudo avaliou os arranjos familiares em função de gênero e por classe de renda, dados demográficos relativos aos núcleos familiares (arranjos familiares), as posições de responsabilidade dos gêneros na condução e manutenção das famílias e os fatores que contribuem para a vulnerabilidade familiar, constatando que os arranjos familiares chefiados por mulheres sozinhas estão mais fadados ao fracasso em termos de oportunidade, tendo como questão básica da análise: “Qual é o fator que mais contribui para uma família ser vulnerável, se ter crianças ou ser chefiada por mulheres sozinhas, que, por sua inserção menos favorável no mercado de trabalho e oportunidades de emprego mais restritas, acabariam limitadas no desenvolvimento do seu potencial produtivo, com implicações negativas no enfrentamento das suas obrigações familiares?”. Como solução à questão anterior, foi utilizada a técnica *logit* para avaliar se uma família é vulnerável ou não diante dos fatores que caracterizam os arranjos familiares.

Foram gerados coeficientes de proporções de inativos, crianças, idosos, contribuintes da previdência, idosos com aposentadoria, tipos de famílias, rendimentos do chefe e rendimento do cônjuge para as famílias constantes na PNAD (2003), foi utilizado o nível de significância de 10% para testar a máxima verossimilhança, e chegaram às conclusões de que, famílias com crianças e idosos tem maior probabilidade de falência e de que famílias chefiados exclusivamente por

mulheres tem menos chances de falência do que àquelas chefiadas de forma biparental, concluindo também que um aumento na renda do cônjuge reduz a probabilidade de falência em um percentual maior do que o aumento da renda do chefe da família.

De acordo com Abbad, Carvalho e Zerbini (2006), apesar da importância dos cursos à distância como ferramenta de desenvolvimento de competências humanas para o trabalho, em análises da literatura nacional e estrangeiras, os sistemas de treinamento à distância, não utilizam análises sistemáticas das necessidades de treinamento, sem relacioná-las às competências críticas organizacionais, fazem pequeno esforço na avaliação da aprendizagem e investem muito pouco no planejamento instrucional de cursos presenciais, além de utilizarem estratégias e meios incompatíveis com o perfil da clientela em modelos presenciais, privilegiando assim, as pessoas lotadas mais próximas aos centros onde são desenvolvidos os treinamentos, em detrimento de modelos facilitadores da transferência horizontal e vertical de conhecimentos e novas aprendizagens. Afirmam também que é pequena a quantidade de pesquisas que avaliam esses cursos e, em particular, os índices de evasão.

O objetivo do trabalho é identificar variáveis explicativas da evasão em um curso gratuito a distância, via Internet, oferecido em nível nacional. Evasão refere-se à desistência definitiva do aluno em qualquer etapa do curso. As variáveis antecedentes incluíram dados demográficos e de uso dos recursos eletrônicos. Foi adotada a análise de regressão logística pela natureza categórica da variável critério (concluinte *versus* não concluinte). Os resultados mostram que os alunos não concluintes são aqueles que tendem a não utilizar os recursos eletrônicos de interação (mural de notícias, *chats*, troca de mensagens eletrônicas). São discutidas

as implicações teóricas e metodológicas dos achados, a necessidade de inclusão de variáveis motivacionais, cognitivas e contextuais, bem como suas possíveis interações em modelos multivariados explicativos de evasão.

Ainda segundo as autoras, há pouco esforço dos pesquisadores em avaliar as ações de treinamento à distância em educação (TD&E), mediadas por novas tecnologias. Segundo as autoras, em pesquisas desenvolvidas na Ásia por Xenos et al. (2002) apud Abbad et al (2006), os índices de evasão nos cursos à distância podem atingir 50% dos inscritos, enquanto que no estudo feito na Europa por Shin e Kim (1999) apud Abbad et al (2006), aponta que o índice de evasão encontra-se entre 20% e 30% dos alunos matriculados.

O estudo conduzido por Abbad et al (2006), foi realizado com a finalidade de identificar variáveis explicativas dos índices de evasão em um curso técnico à distância oferecido via internet, em termos de características da clientela e da interação dos participantes com recursos eletrônicos disponibilizados pelo curso. Foi aplicado um modelo utilizando a Regressão Logística binária, visando identificar, entre algumas variáveis da clientela do curso, quais explicavam os índices de evasão.

Segundo Xenos et al. (2002) apud Abbad et. al (2006), os níveis de evasão em cursos universitários à distância, são afetados por fatores internos relacionados às percepções dos alunos e seu locus de controle (interno e externo); fatores relacionados ao curso e aos tutores; e fatores relacionados a certas características demográficas dos estudantes, como idade, sexo, estado civil, número de filhos, tipo de trabalho ou profissão, entre outras.

Em estudo realizado por Shin e Kim (1999) apud Abbad et al. (2006), na Universidade Nacional Aberta da Coreia, definiram variáveis que influenciam na

evasão em cursos à distância (endógenas e exógenas). As variáveis exógenas foram identificadas como *carga de trabalho* (percepção do participante sobre o grau de exigência do trabalho que executa em seu emprego), *integração social* (medida em termos das percepções que o participante possui sobre o apoio das pessoas que o rodeiam para estudar) e *anseio* (o desejo do aluno em concluir o curso). Foram identificadas como variáveis endógenas, o tempo de estudo, o planejamento da aprendizagem e atividades face a face. No estudo de Shin e Kim (1999) foi realizada uma análise de regressão logística para investigar os fatores associados à evasão de alunos.

A pesquisa de Abbad et al. (2006) é de natureza exploratória e busca correlações e de campo, foi realizada no período de janeiro a dezembro de 2002. Foi conduzida a partir a avaliação de um curso gratuito, realizado via internet, totalmente à distância, com tutoria ativa e que visava a ensinar a clientela a elaboração de um plano de negócios. O modelo de investigação propõe-se a avaliar as características demográficas da clientela, o uso dos recursos eletrônicos do curso e a evasão dos alunos.

Características da Clientela refere-se aos dados demográficos dos participantes, como sexo, idade, região geográfica e nível de escolaridade. O uso dos recursos eletrônicos do curso diz respeito à frequência de utilização espontânea das ferramentas: *chats*, listas de discussão e ambiente eletrônico do curso (*website* e *webpages*) durante todo o período do curso. O terceiro componente, evasão, variável critério neste estudo, refere-se à desistência definitiva do aluno em qualquer etapa do curso, medida em termos da quantidade de alunos concluintes e não concluintes do curso.

De modo a avaliar o efeito do tamanho elevado da população (N = 19.849) sobre a significância das correlações entre variáveis e a consistência dos resultados (erro tipo I), Tabachnick e Fidell (2001) sugerem que se realizem as mesmas análises com amostras aleatórias extraídas da população estudada.

Sendo assim, nesta pesquisa foram repetidas as análises de regressão logística em duas amostras extraídas da população (amostra 1: n = 1.349; amostra 2: n = 1.436). As amostras são aleatórias simples e foram geradas pelo próprio programa de análise estatística.

Ainda segundo Abbad et al. (2006), foram realizadas análises descritivas e exploratórias para investigar a exatidão da entrada dos dados, a presença de casos extremos, a distribuição dos casos omissos e as características de distribuição de frequência das variáveis. Seguindo orientação de Tabachnick e Fidell (2001), as autoras optaram pela análise de regressão logística binária em função da natureza categórica da variável critério (*concluente versus não concluente*).

O primeiro conjunto de variáveis antecedentes corresponde a variáveis demográficas da clientela do curso. São elas: sexo, idade, nível de instrução e região geográfica de residência.

A análise logística para a população de inscritos resultou na obtenção das correlações entre as variáveis, dos coeficientes de regressão não padronizados, o intercepto (constante) e do erro padrão, bem como o coeficiente Wald, as estatísticas Cox & Snell R², Nagelkerke R² e a probabilidade dos resultados observados (-2LL).

O teste do modelo completo com as dez variáveis antecedentes apresentou uma probabilidade dos resultados observados (-2LL) igual a 16.696,40, sendo este valor relativo ao *log likelihood value*. Para se avaliar o poder preditivo das variáveis

anteriores, analisam-se as estatísticas Cox & Snell R^2 e Nagelkerke R^2 , as quais são semelhantes ao R^2 do modelo de regressão múltipla linear, mais comumente utilizado em pesquisas. Portanto, pode-se dizer que 55,6% da variação no resultado da variável dependente (VD) é explicada pelo modelo de regressão logística. Isso significa que, em mais da metade dos casos, a evasão (VD) é explicada pelo modelo testado neste estudo.

O teste do modelo completo, comparado com o modelo da constante, apresentou diferença significativa, indicando que oito das 10 variáveis anteriores diferem significativamente entre concluintes e não concluintes ($p < 0,01$). Apenas as variáveis sexo e região geográfica não apresentam diferença significativa, apesar de esta última apresentar diferença a 5% de significância. Outra maneira de avaliar o ajuste do modelo consiste na comparação entre os valores preditos e os resultados realmente observados (teste qui-quadrado). O ajuste do modelo foi expressivo, com 87,8% de concluintes e 75,1% de não concluintes classificados corretamente. No total, o modelo classificou corretamente 82% de participantes.

Para Zeilhofer et al. (2007), as relações causais entre uso agrícola na qualidade de água foram comprovadas em diversos estudos, entre eles os de Fisher et al. (1999) apud Zeilhofer et al. (2007) e Stow et al. (2001) apud Zeilhofer et al. (2007). Em países com intensa atividade agrícola tais como os Estados Unidos, a poluição difusa ou não-pontual é considerada o maior problema para a qualidade de água, introduzindo nos ecossistemas aquáticos sobrecargas de nutrientes, pesticidas e sedimentos.

Como base de estudos sobre impactos e origens de problemas de qualidade de água, técnicas de sensoriamento remoto e geoprocessamento podem ser

aplicados no mapeamento de uso e ocupação de solo e na parametrização das bacias de contribuição.

Além de uma variável não-especializada “Período sazonal”, foram avaliadas as supostas relações com os PIs do “Uso e ocupação de solo”, das “unidades pedológicas”, “aptidão agrícola”, “unidades geomorfológicas”, “declividade” e “distância para o manancial mais próximo”. A variável “distância para o manancial mais próximo” foi logaritmizada para homogeneização dos dados.

As principais saídas do modelo de Regressão Logística (*backward conditional*), que indicou o período sazonal (“Chuva”) e a “Distância para os mananciais” (“Distancia”) como fatores explanatórios significativos. O resultado foi um modelo com nível de acerto de 86,2%.

Segundo CASCARINO (2007), em complemento à utilização de métodos estatísticos, há vários métodos quantitativos à disposição dos auditores internos na condução dos trabalhos de auditoria. Estas ferramentas matemáticas são normalmente utilizadas para analisar e entender as operações realizadas pelos sistemas, permitindo a elaboração de conclusões sobre uma variedade de situações de grande complexidade que são objeto das análises, dentre estas, encontram-se a análise de tendências, testes qui-quadrado, análise de correlação, análise de regressão linear e de probabilidade.

A análise de tendência é utilizada para avaliar o comportamento da mudança de uma variável (Variável dependente) ao longo do tempo, servindo para identificar o critério e determinar o motivo de variações (Variáveis independentes) em um determinado período (passado, presente ou futuro).

O teste qui-quadrado é um teste não-paramétrico e possibilita a análise de relacionamentos entre dados qualitativos e permite analisar as distribuições normais

e dados ordinais, não necessitando de distribuições específicas do conteúdo. A correlação é a medida que quantifica o nível de associação entre as variáveis comparadas, sendo correlacionadas quando há um padrão de variação de incremento ou decréscimo.

Os auditores internos utilizam-se de análises gráficas para identificar anomalias, associações e erros nos dados. Uma forma de representação gráfica normalmente utilizada é o *scatter diagram*, estes reportam os vários comportamentos e tendências de todos os pontos do gráfico, suas correlações e a possibilidade de prever valores, baseando-se no comportamento da série de dados apresentada.

As correlações só conseguem demonstrar o comportamento de padrões lineares, portanto, padrões não lineares podem distorcer a análise gráfica do *scatter plot*, sendo necessário utilizar análises de *Kolmogorov-Smirnov* para verificar se a distribuição dos valores segue ou não os padrões normais de distribuição. Na condução de avaliações dos níveis de desempenho relativos à implementação de novos procedimentos de qualidade, a curva de aprendizado normalmente é utilizada, comparando o desempenho esperado e o observado, possibilitando inferir sobre os tempos necessários à realização de atividades de produção e desenvolvimento das atividades organizacionais.

Quando novos procedimentos são recomendados, devem ser calculados os tempos por unidade de produção na metodologia anterior, possibilitando a comparação com os novos tempos utilizados, permitindo a comparação entre os tempos e aferindo o impacto da implantação do procedimento nos processos produtivos.

A análise das razões consiste em obter as proporções dos relacionamentos entre dois números ao longo do tempo. Uma forma mais aprimorada da análise das razões utilizada para mensurar o relacionamento entre variáveis é a análise de regressão, utilizada para avaliar o efeito de um conjunto de variáveis (variáveis independentes) sobre outra variável alvo do estudo (variável dependente), possibilitando a predição de valores da variável dependente a partir de um conjunto (série) de observações das variáveis independentes ao longo do tempo.

Nas aplicações de auditoria, as análises de regressão auxiliam o auditor a entender e quantificar o relacionamento das informações, permitindo a detecção de variações inesperadas entre valores e variáveis, levando à avaliações mais criteriosas sobre estas observações.

Através da utilização de software, o auditor poderá executar regressões buscando discriminar as interferências das variáveis independentes sobre a variável dependente, determinando a intensidade do relacionamento, direcionando ações para evitar os impactos destes comportamentos sobre as atividades das organizações.

2.3 Indicadores Antecedentes na área social

Para Sampaio et al. (1999), no Brasil os acidentes de trabalho (AT), são legalmente definidos como: típico (quando ocorrem durante a jornada de trabalho), de trajeto (quando ocorrem entre a residência e o local de trabalho) e doença-profissional (quando decorrem do exercício profissional), segundo a Lei no. 8.213 de 24/02/1991 (MPAS, 1995b). Este trabalho avaliou o término da incapacidade através de um estudo transversal dos trabalhadores incapacitados para o exercício de suas profissões habitais, submetidos ao programa CRP/INSS de Belo Horizonte, Minas

Gerais, Brasil. Foram utilizados dados secundários extraídos dos históricos médicos a partir do ingresso do trabalhador no CRP (Centro de Reabilitação Profissional).

O estudo busca a avaliação dos fatores, através da técnica estatística de Regressão Logística, que estão associados à probabilidade de reingresso dos trabalhadores incapacitados permanentemente para o exercício de suas profissões no mercado de trabalho, após re-qualificados para outras atividades através dos Centros de Reabilitação Profissional mantidos pelo Ministério de Previdência e Assistência Social (MPAS).

Foram estudadas as variáveis demográficas sexo, idade, estado civil, escolaridade, motivo do ingresso no programa, informações sobre o tipo de trabalho executado, ocupação segundo o CBO (Cadastro Brasileiro de Ocupações), setor econômico (de acordo com o CNAE – Cadastro Nacional de Atividades Econômicas), situação do trabalhador (autônomo, assalariado, desempregado, etc..), diagnóstico médico e o resultado da atuação do programa de reabilitação. Foram analisados 920 históricos médicos, feitas análises descritivas utilizando o software estatístico, chegando aos resultados de que 67% dos casos tratavam-se de acidentes por doenças profissionais, 31% referiam-se aos acidentes típicos e 2% eram relativos aos acidentes de trajeto.

Como resultado, foi identificado que 58% dos casos eram constituídos por homens e 42% de mulheres, a média de idade encontrada foi de 34 anos. Em relação ao estado civil, 63% dos participantes eram casados, 31% solteiros, 5% separados e 1% viúvos. Em relação ao nível de instrução, 50% dos trabalhadores tinham 1º grau incompleto e a maioria deles (90%) eram assalariados e encontravam-se formalmente inseridos no mercado de trabalho.

Em relação às atividades desempenhadas segundo o CBO, 9,4% dos trabalhadores eram auxiliares de escritório, 8,5% eram empregados em bancos e auxiliares contábeis, 7,7% atuavam em serviços de limpeza e manutenção e finalmente, 6,4% dos trabalhadores atuavam em equipamentos de processamento de dados. Em relação ao CNAE, os trabalhadores encontram-se distribuídos pelas instituições financeiras, seguidos pelo setor de construção civil, metalúrgica e prestadores de serviços, estes distribuídos pelas atividades de temporários, limpeza e vigilância.

Em relação ao tempo de trabalho, o padrão é diferente para os casos de acidente e doenças-profissionais. Para os casos de acidente, quase a metade levava menos de 6 meses trabalhando, enquanto que para as doenças-profissionais, 26% dos casos estavam entre 4 e 9 anos de trabalho, seguidos de 25% dos trabalhadores que teriam entre 6 meses a 2 anos de trabalho.

Mais da metade dos casos analisados (53%) foram diagnosticados como LER (Lesão por Esforço Repetitivo), seguidos por amputações traumáticas dos dedos das mãos (3,4%), efeitos tardios de lesões traumáticas ósteo-musculares (3,3%), dermatites de contato (3,2%) e doenças pulmonares em decorrência de partículas de sílica.

Dos 920 casos analisados pelo estudo de Sampaio et al. (1999), 35% (322) foram aceitos pelo programa de reabilitação após avaliados pelos técnicos do centro. Entre os 592 casos avaliados e aceitos para o programa de reabilitação profissional, 25% (142 trabalhadores) foram considerados sem possibilidade de voltar ao trabalho depois de iniciada a reabilitação profissional.

O resultado do modelo de regressão logística apontou que o sexo dos trabalhos tem pouca ou nenhuma interferência na possibilidade de ser aceito no

programa, enquanto que foram identificados como os que possuem maior chance de aceitação no programa, aqueles trabalhadores que sejam portadores de LER, com menos de 40 anos e com nível médio de escolaridade (primeiro grau completo ou incompleto e segundo grau incompleto). Os trabalhadores que atendam a estas condições possuem 75,6% de chances de serem aceitos, segundo o modelo logit gerado. A presença ou não de LER não influencia no resultado do programa.

O modelo *logit* apresentou a probabilidade de 90,9% de retorno ao mercado de trabalho formal para os trabalhadores homens, com menos de 40 anos, com o mais alto nível de escolaridade (segundo grau completo ou superior incompleto).

3 METODOLOGIA

A metodologia utilizada no trabalho foi aplicada em nove fases, a saber:

1. Realização de uma enquete com os gerentes e coordenadores que atuam na área de benefícios por incapacidade da DATAPREV e Previdência Social, com o propósito de identificar as variáveis que deverão compor o escopo da análise;
2. Realização de uma pesquisa exploratória sobre as bases de dados da Previdência Social buscando obter informações quantitativas sobre as variáveis identificadas na enquete realizada previamente, verificando a relevância das mesmas e os níveis de significância no universo pesquisado (Benefícios por incapacidade atendidos e gerenciados pela Previdência Social);
3. Utilizar um questionário auto-aplicado no âmbito da Previdência Social e DATAPREV, visando à obtenção de manifestações do corpo funcional envolvido diretamente nas atividades de manutenção e concessão de benefícios por incapacidade, com o propósito de obter informações detalhadas sobre o nível de influência das variáveis para a categorização de critérios do perfil dos benefícios cessados, possibilitando a seleção das variáveis que serão verificadas no estudo corrente;

A partir das informações relativas aos benefícios por incapacidade oriundos das bases de dados da Previdência Social, gerar um modelo que seja capaz de apoiar as verificações sobre as variáveis identificadas no processamento dos questionários aplicados na fase 03 (três) desta proposta metodológica;

4. Promover a seleção dos benefícios que encontram-se na categoria de “cessados” em função de incorreções identificadas pela Auditoria da Previdência Social ou pela Inspeção da Previdência Social;
5. Aplicação do modelo de Regressão Logística sobre as informações identificadas no item 4 e obtenção da equação logística sobre estes benefícios;
6. Aplicação da equação obtida no item 5 (cinco) sobre os benefícios por incapacidade constantes nas bases de dados como “ativos”;
7. Apresentação dos resultados obtidos sobre a aplicação dos padrões obtidos a partir de benefícios “cessados” sobre os benefícios da mesma categoria, gerando *scores* de chance de irregularidade para estes benefícios que encontram-se na condição de “ativos”, sendo pagos regularmente;
8. Discussão dos resultados obtidos a partir da aplicação do modelo de Regressão Logística sobre as categorias de benefícios por incapacidade, considerando o *score* de chance de irregularidade e indicando um modelo para integrar as ações de verificação de conformidade dos benefícios por incapacidade em futuras ações promovidas pela Previdência Social através de seus órgãos de controle e monitoramento.

Segundo Demo (1985, p.19) :

“A metodologia é uma preocupação instrumental. Trata das formas de fazer ciência. Cuida dos procedimentos, das ferramentas, dos caminhos. A finalidade da ciência é tratar a realidade teórica e praticamente.”

Para Matias-Pereira (2007), “a função social da pesquisa está implícita na própria visão de mundo e de ciência do pesquisador”, afirmando ainda que a motivação do pesquisador é o fator impulsionador da prática científica, alegando

também que, a motivação do pesquisador “dará conta dos caminhos escolhidos como curso investigativo”.

Para Barros e Lehfeld (2000, p.78) apud Matias-Pereira (2007), há relações entre a pesquisa teórica e a pesquisa aplicada, e:

“... enquanto na pesquisa teórica o pesquisador está voltado a satisfazer uma necessidade intelectual de conhecer e compreender determinados fenômenos, na pesquisa aplicada ele busca orientação prática à solução imediata de problemas concretos do cotidiano.”

Para Gil (1999, p. 20), “a ciência pode ser considerada como uma forma de conhecimento objetivo, racional, sistemático, geral, verificável e falível”. O conhecimento é objetivo porque descreve a realidade, independentemente de interferências ou interpretações do pesquisador. É racional em função da sua fundamentação na razão e não em sensações ou impressões para a obtenção dos resultados. É sistemático em função da preocupação em construir sistemas de idéias organizadas racionalmente, visando ampliar os conhecimentos parciais em esferas maiores. O conhecimento é geral porque destina-se à elaboração de leis ou normas gerais que explicam os fenômenos. É verificável porque sempre possibilita demonstrar a veracidade das informações analisadas, sendo também falível e reconhecendo a sua capacidade de errar.

Segundo Yin (1981) apud Roesh (1995, p.146), o estudo de caso é uma estratégia de pesquisa que busca examinar um fenômeno dentro do seu contexto, diferindo de métodos históricos por referirem-se ao presente e não ao passado. Diferem também de delineamentos experimentais por analisarem os fenômenos estudados em seu próprio contexto.

Em Roesh (1995, p.147), afirma-se que o estudo de caso como uma estratégia de pesquisa pode ser “utilizado de modo exploratório (visando levantar

questões e hipóteses para futuros estudos), descritivos (buscando associações entre variáveis)”.

Para Gil (1999, p. 72), o estudo de caso é caracterizado pelo “estudo profundo de um ou poucos objetos de maneira a permitir o seu conhecimento amplo e detalhado”, sendo esta tarefa “praticamente impossível mediante os outros tipos de delineamento considerados”.

De acordo com Yin (1984, p. 23), o estudo de caso é um estudo empírico que investiga um fenômeno atual “dentro do seu contexto de realidade, quando as fronteiras entre o fenômeno e o contexto não são claramente definidas”.

Segundo Gil (1999, p. 73), o estudo de caso vem sendo utilizado com frequência cada vez maior, atendendo aos propósitos de explorar situações da vida real devido ao fato dos limites não serem claramente definidos, descrever a situação do contexto em que está sendo feita a investigação e explicar as variáveis causais de determinados fenômenos.

Para Gil (1995), o estudo de caso não aceita um roteiro rígido para a sua delimitação, mas constitui-se de quatro fases, designadas como:

a) delimitação da unidade-caso, onde são selecionados casos atípicos, casos extremos e casos típicos;

b) coleta de dados, feita através de levantamentos de dados sobre a unidade-caso selecionada;

c) seleção, análise e interpretação dos dados que deve considerar os objetivos da investigação, seus limites e um sistema de referências para avaliar quais dados serão úteis ou não, sendo somente analisados os dados que foram selecionados;

d) elaboração do relatório, deve apresentar os dados e especificar como foram coletados, a teoria que embasou a categorização dos mesmos e a demonstração da validade e da fidedignidade dos dados obtidos.

Para Goode e Hatt (1979), o estudo de caso é um meio de organizar os dados, preservando do objeto estudado o seu caráter unitário.

Para Ventura (2007), há vantagens decorrentes da aplicação dos estudos de caso, tais como: estímulo às novas descobertas, em função da flexibilidade do seu planejamento; enfatiza a multiplicidade de dimensões de um problema, focalizando-o como um todo e apresenta simplicidade nos procedimentos, além de permitir uma análise em profundidade dos processos e das relações entre eles.

Para Ventura (2007), há limitações neste tipo de estudo:

A mais grave, parece ser a dificuldade de generalização dos resultados obtidos. Pode ocorrer que a unidade escolhida para investigação seja bastante atípica em relação às muitas da sua espécie. Naturalmente, os resultados da pesquisa tornar-se-ão bastante equivocados. Por esta razão, cabe lembrar que, embora o estudo de caso se processe de forma relativamente simples, pode exigir do pesquisador muita atenção e cuidado, principalmente porque ele está profundamente envolvido na investigação.

Neste trabalho foi apresentada uma Pesquisa Aplicada de Campo de natureza quantitativa, através da aplicação da técnica estatística inferencial de Regressão Logística sobre os benefícios relativos às Aposentadorias por Invalidez e Auxílios por Acidente do Trabalho (*Benefícios por incapacidade*), nas categorias de *cessados* ou *suspensos* por determinação da Auditoria ou pela Inspeção do MPAS, Órgãos de Controle da Previdência Social Brasileira. As informações foram obtidas a partir das bases de informações da Previdência Social na competência dezembro/2007.

A Pesquisa Aplicada tem como objetivo gerar conhecimentos para a aplicação prática, estes conhecimentos serão direcionados às soluções de problemas

específicos, envolvendo verdades e interesses locais (MATIAS-PEREIRA, 2007, p. 70). Já para Michel (2005, p. 36), a pesquisa aplicada tem sua utilização através dos resultados adquiridos na pesquisa básica, voltando-se mais para o aspecto utilitário da pesquisa, procurando transformar o “conhecimento puro em elementos, situações destinadas a melhorar a qualidade de vida da humanidade”.

Para Lakatos e Marconi (2007b, p. 20), a pesquisa aplicada caracteriza-se por seu interesse prático, isto é, que a aplicação dos resultados seja feita imediatamente na solução de problemas que ocorrem na realidade.

Na pesquisa quantitativa o pesquisador descreve, explica e prediz, utiliza também a quantificação tanto na coleta de dados quanto no tratamento das informações através de técnicas estatísticas, desde a mais simples (percentuais, média, desvio-padrão) até as mais complexas (correlação, análise de regressão), a utilização destas técnicas garante a intenção, neste tipo de pesquisa, de assegurar a precisão dos resultados, evitando distorções de análise e interpretação, garantindo a margem de segurança quanto às inferências (MICHEL, 2005, p. 33). Ainda segundo a autora:

A pesquisa quantitativa se realiza na busca de resultados precisos, exatos, comprovados através de medidas de variáveis preestabelecidas, na qual se procura verificar e explicar sua influência sobre outras variáveis, através da análise da freqüência de incidências e correlações estatísticas. Os resultados são obtidos e comprovados pelo número de vezes em que o fenômeno ocorre ou com a exatidão em que ocorre. A resposta que se busca na investigação deve ser obtida de forma numérica, exata, inquestionável.

Segundo Rodrigues (2007, p.42), a pesquisa de campo busca fontes primárias “no mundo dos acontecimentos não provocados e nem controlados pelo pesquisador, que se caracteriza por desenrolar-se em ambiente natural”. Utiliza o procedimento de observação direta do objeto estudado no meio em que é próprio, sem a interferência do pesquisador.

Foram buscadas as identificações das variáveis preditoras das naturezas demográfica, geográfica e sócio-econômica, capazes de detectar o comportamento dos benefícios nas categorias de Acidentes do Trabalho e Aposentadorias por Invalidez (*Benefícios por incapacidade*), identificando em função do nível de conformidade como “Corretos” ou “Incorretos” os benefícios que estiveram enquadrados nos critérios resultantes da análise inferencial multivariada.

O modelo resultante foi aplicado sobre o total de benefícios na condição de *ativos* em dezembro de 2007 com o propósito de identificar o grau de conformidade e adequação à legislação vigente dos mesmos, possibilitando que o INSS realize testes detalhados, verificando inclusive o nível de predição do modelo sobre o universo dos *benefícios por incapacidade* que são pagos pela Previdência Social, buscando a suspensão dos incorretos e a recuperação de possíveis valores pagos indevidamente.

3.1 Método Indutivo

Para Marques et al. (2006), “...na indução parte-se da amostra (concreto) para o abstrato (geral) na tentativa da generalização...”, indo do particular para o geral, sendo as constatações particulares que levam às teorias e leis gerais. Ainda segundo Marques e al. (2006):

A conclusão estabelecida pela indução pode servir de princípio – premissa maior – para a dedução, mas a conclusão da dedução pode também servir de princípio da indução seguinte – premissa menor –, e assim sucessivamente.

O método Indutivo proposto pelos empiristas Bacon, Hobbes, Locke e Hume, considera que o conhecimento é fundamentado na experiência, não levando em conta princípios preestabelecidos. No raciocínio indutivo a generalização deriva de

observações de casos da realidade concreta. As constatações particulares levam à elaboração de generalizações (GIL, 1999; LAKATOS e MARCONI, 2008).

Segundo Michel (2005), o silogismo é uma forma de raciocínio, um método para chegar-se à conclusão, baseado em três componentes: premissa maior; premissa menor e conclusão. A premissa maior deverá conter a verdade em seu limite máximo, ser “universal” dentro do propósito estabelecido. A premissa menor deverá estar relacionada com a verdade da premissa maior, podendo ou não ser comprovada integralmente por ela.

O quadro 2 apresenta o resumo teórico do estudo, identificando a tipologia, tipo de estudo, tipo de argumento e o método utilizado na coleta de dados.

Quadro 2 - Resumo teórico do estudo

Tipologia da Pesquisa - Pesquisa Aplicada de Campo e Estudo de Caso
Abordagem – Quantitativa
Tipo de Estudo - Descritivo
Tipo de Argumento - Indutivo
Método de coleta de dados - Levantamento de dados

Fonte: Autor do estudo

Em relação à amostragem Cooper e Schindler (2003, p.80), definem que o pesquisador deve determinar quais e quantos eventos observar e quais e quantos registros inspecionar. Uma amostra é parte de uma população-alvo e deve representar esta população. Afirmam ainda que “quando os pesquisadores conduzem estudos por amostragem, estão interessados em estimar um ou mais valores da população e em testar uma ou mais hipóteses estatísticas”. Afirmam também que “o processo de amostragem deve dar a cada pessoa dentro da população-alvo uma chance diferente de zero para seleção, se a amostragem for probabilística”.

Segundo Cooper e Schindler (2003, p. 167), a utilização da amostragem não-probabilística possui a capacidade desconhecida de selecionar elementos da população.

Em Cooper e Schindler (2003, p. 84), considera-se que a análise de dados envolve a redução dos dados acumulados a um tamanho administrável, o desenvolvimento de sumários, a busca de padrões e a aplicação de técnicas estatísticas. As respostas da pesquisa exigem que o analista “derive várias funções e também explore as relações entre as variáveis”, buscando a verificação de que os resultados são consistentes com as hipóteses e teorias.

De acordo com Malhotra (2006, p. 101), há a possibilidade do uso da pesquisa descritiva quando os objetivos forem descrever as características de grupos relevantes, como consumidores. O autor também indica a pesquisa descritiva quando o objetivo for estimar a porcentagem de pessoas em uma população específica que exibe um determinado comportamento, ou ainda, para determinar as percepções de características.

3.2 Survey de experiência

Para Cooper e Schindler (2003), apenas parte do conhecimento existente em uma área organizacional está documentado e quando está escrito, pode ser confidencial. A partir deste pressuposto, Cooper e Schindler (2003) alegam que:

“[...] é muito útil buscar informações com pessoas experientes na área de estudo, extraindo essas informações de suas memórias e experiências coletivas.”.

Segundo Cooper et Schindler (2003), o *Survey de experiência* é uma entrevista onde devem ser obtidas as idéias dos especialistas em relação às questões ou aspectos importantes do assunto tratado e descobrir o que é importante

dentro do campo de conhecimento da pessoa. O formato investigativo que será utilizado deve ser flexível para explorar as várias possibilidades que surgirem durante a entrevista.

Na busca por referencial teórico sobre as características relativas aos benefícios por incapacidade para nortear o estudo, chegou-se à constatação de sua inexistência no âmbito da Previdência Social, neste estudo foram considerados trabalhos das Auditorias do Ministério da Previdência e Assistência Social e DATAPREV, sendo tomada a decisão de promover um *survey de experiência* junto aos técnicos, gerentes e especialistas da DATAPREV e do MPAS que atuam nas atividades de manutenção e concessão destas categorias de benefícios, buscando a partir das experiências destes, a identificação de um conjunto de características dos benefícios por Incapacidade, visando a composição de um conjunto de variáveis passíveis de análise no estudo.

3.3 Definição das variáveis

De acordo com Boaventura (2007), as variáveis constituem um dos principais fundamentos da ciência. Afirma também, que uma variável é algo que muda de valor, tornando-se diferente e que variáveis que possuam apenas um valor são designadas como constantes.

Segundo Kerlinger (1980) apud Boaventura (2007), as variáveis são propriedades que podem assumir valores diferentes. Afirma também que as variáveis podem assumir muitos valores, porém estas devem assumir no mínimo dois valores.

Segundo Cooper et Schindler (2003, p. 57), um traço é definido como uma “...entidade ou coisa sobre a qual a hipótese discorre.”, enquanto que uma variável “...é uma característica, o traço ou o atributo que na hipótese é imputada ao caso.”.

Para Lakatos et Marconi (2008), a variável pode ser considerada uma classificação ou ainda uma medida, uma quantidade que varia, um conceito, um constructo que contém ou apresenta valores. Os valores que são acrescentados ao conceito podem ser quantidades, qualidades, características, magnitudes ou traços, que se alteram em cada caso particular. As autoras afirmam ainda que uma nova aplicação para a definição operacional é a transformação de constructos ou conceitos em variáveis, que caracterizam-se pela possibilidade de mensurá-las, em função de conterem valores ou classe de valores.

Ainda segundo Lakatos et Marconi (2008, p. 175):

[...] podemos imaginar o “universo” da ciência como constituído de três níveis: no primeiro, ocorrem as observações de fatos, fenômenos, comportamentos e atividades reais; no segundo, encontramos os conceitos ou constructos em forma de hipóteses; finalmente, no terceiro, surgem as teorias, hipóteses válidas e sustentáveis, compostas de constructos e termos teóricos. O que nos interessa, na realidade, é a passagem do segundo para o primeiro nível, o que ocorre mediante a definição operacional, com a delimitação das variáveis.

Após a avaliação das variáveis identificadas a partir da análise dos questionários respondidos pelos especialistas da DATAPREV e do INSS, chegou-se às informações retratadas na tabela 4, onde observa-se a forte ênfase na análise de aspectos relativos às regras da Previdência Social em detrimento de aspectos demográficos relativos aos segurados.

Tabela 4 - Definição da origem das variáveis

Tipos de variáveis	Quantidade
Demográficas segurados	3
Demográficas respondentes	3
Previdenciárias	17
Total	23

Fonte: Elaboração própria

3.4 Definição do questionário

Segundo Aaker et al. (2001), em relação aos questionários, alega que a maioria dos estudiosos concorda que os pontos da escala Likert devem ter mais de três alternativas de respostas, porém, não deve ultrapassar nove.

Nas pesquisas de opinião, o interesse preponderante é identificar as relações entre as variáveis, o que requer variância nas respostas em relação às questões formuladas. Expandir o número de respostas quase sempre aumenta a variância nas respostas (BABBIE, 2001).

O número adequado de opções de respostas às questões constantes no questionário, depende do grau de conhecimento dos respondentes sobre o assunto que está sendo tratado no mesmo. Para especialistas na área de pesquisa em foco, a escala deve contemplar um maior número de opções de respostas (PEARSON, 2003).

Neste estudo, os questionários serão utilizados para a obtenção de dados primários sobre a relação entre as variáveis pesquisadas e o comportamento e conformidade dos benefícios por Incapacidade devido à inexistência de referenciais, no âmbito da Previdência Social, sobre as questões pesquisadas.

Para Cavusgil et Elvey-Kirk (1998), o pesquisador deve assegurar o anonimato do respondente, fortalecendo o fator motivador do respondente em relação aos resultados sociais positivos.

O questionário foi enviado através de e-mail corporativo (somente no âmbito da Previdência Social e DATAPREV), acompanhado de uma carta onde foram identificadas a relevância das informações prestadas, a utilização das mesmas e a sua aplicação em benefício da sociedade, demonstrando assim, a importância da

expressão da opinião do respondente e o retorno do questionário no menor prazo possível.

O questionário é um instrumento de coleta de dados, construído por uma série ordenada de perguntas, que devem ser respondidas por escrito, sem a presença do entrevistador. Junto com o questionário deve ser enviada uma nota ou carta explicando a natureza da pesquisa, sua importância e a necessidade de obter as respostas. Tais procedimentos visam sensibilizar o respondente para que este preencha e devolva o questionário em um prazo razoável (LAKATOS et MARCONI, 2007b, p. 98).

O questionário como mecanismo de coleta de dados apresenta uma série de vantagens em relação às outras formas de coleta, tais como: economia de tempo, viagens e obtenção de uma grande quantidade de informações; atinge um maior número de respondentes, inclusive em áreas geográficas diferentes; obtém respostas mais rápidas e mais precisas devido à preservação do anonimato do respondente; há menor risco de distorção em função do pesquisador não influenciar nas respostas; há mais tempo para responder e fornecer as respostas, possibilitando o preenchimento em horário favorável; a natureza impessoal do instrumento possibilita uma maior uniformidade na avaliação das informações fornecidas pelos respondentes (LAKATOS et MARCONI, 2007b).

Lakatos et Marconi (2007b), alertam para alguns problemas que ocorrem quando o questionário é utilizado como instrumento de coleta de dados, tais como: a reduzida quantidade de questionários retornados; uma grande quantidade de questões retornadas sem respostas; não pode ser aplicado a pessoas analfabetas; a devolução tardia prejudica o calendário ou a sua utilização e a aplicação do instrumento exige um universo mais homogêneo.

Ainda segundo Lakatos et Marconi (2007b), o questionário deve conter o resultado de uma avaliação criteriosa sobre o tema, facilitando a percepção dos respondentes sobre as questões, ser limitado em extensão e finalidade, possuir de 20 a 30 questões e demorar em torno de 30 minutos para ser preenchido pelos informantes. As autoras definem ainda que o processo de pré-teste (aplicação do questionário em uma amostra do universo-alvo do instrumento de coleta), deverá considerar se o questionário apresenta três importantes elementos: Fidedignidade – Qualquer pessoa que o aplique obterá sempre o mesmo resultado; Validade – Os dados recolhidos são necessários à pesquisa e Operatividade – Vocabulário acessível e significado claro.

Algumas das principais vantagens de um questionário é que nem sempre é necessária a presença do pesquisador para que o informante responda as questões. Além disso, o questionário consegue atingir várias pessoas ao mesmo tempo obtendo um grande número de dados, podendo abranger uma área geográfica mais ampla se for este o objetivo da pesquisa. Ele garante também uma maior liberdade das respostas em razão do anonimato, evitando vieses potenciais do entrevistador. Geralmente, através do questionário, obtêm-se respostas rápidas e precisas (LAKATOS et MARCONI, 2007b).

Mesmo sofrendo muitas críticas, o questionário continua sendo muito utilizado nas diversas áreas. Algumas desvantagens da sua utilização são: a percentagem de retorno dos questionários enviados pelo correio geralmente é pequena e quando a devolução é tardia prejudica o andamento da pesquisa. Muitas vezes há um número grande de perguntas sem respostas. Outra desvantagem é a dificuldade de compreensão da pergunta por parte do respondente quando o pesquisador está ausente.

Foi aplicado um questionário contendo 23 questões estruturadas cujo objetivo principal foi buscar a confirmação das variáveis que influenciam os benefícios por Incapacidade, na opinião dos respondentes, que serão constituídos de especialistas na gestão e operação de tecnologia da informação e gestão previdenciária, sobre benefícios por incapacidade da Previdência Social e DATAPREV. O questionário foi estruturado através da escala Likert de sete níveis.

Foram enviados 200 (duzentos) questionários aos respondentes, via e-mail, sendo o público-alvo desta pesquisa composto de pessoal atuante na área de Seguro Social em níveis de Gestão e Controle dos processos e Tecnologia da Informação, bem como funcionários da Previdência Social envolvidos em atividades correlatas no atendimento e gestão das atividades acidentárias.

A seleção dos respondentes seguiu o critério de uma Amostragem Probabilística Estratificada. Segundo Lakatos et Marconi (2007b, p. 48), “os estratos são formados pelo pesquisador, segundo as necessidades do seu estudo”. Ainda segundo Lakatos e Marconi (2007b), devem ser observadas algumas restrições, tais como, o cuidado para que um mesmo indivíduo não esteja simultaneamente em mais de um estrato em função do atributo utilizado para a formação do estrato. Neste estudo os estratos foram definidos a partir do nível de atuação de cada indivíduo nas atividades sobre benefícios por Incapacidade (cargo, função, atividade).

A efetiva quantificação dos estratos, foi associada à quantidade de questionários considerados (depois de preenchidos e retornados), uma vez que a quantidade de respondentes que não devolveu os questionários, para Lakatos et Marconi (2007b), constitui as não-respostas, ou ainda questões incompletas ou as rasuras e falhas no preenchimento. Somente foram contabilizadas e consideradas

após o retorno, situação a partir da qual este estudo pode inferir sobre o comportamento dos estratos em termos de eficiência.

Foi feito um pré-teste com 10 respondentes, e observamos que o tempo médio para concluir as respostas ao questionário encontrava-se entre 10 e 15 minutos, tempo este considerado aceito segundo Lakatos et Marconi (2007b), que citam como tempo válido o intervalo de 30 minutos para que os respondentes concluam o preenchimento das respostas contidas nos questionários.

O questionário que foi utilizado na obtenção das opiniões dos especialistas consta no APÊNDICE A e as descrições relativas às questões que dependem de identificação pelo respondente sobre mais de uma categoria encontram-se disponíveis no APÊNDICE B.

3.5 Processamento dos questionários

No processamento dos questionários foi utilizada a técnica de consenso, entre os respondentes, de acordo com Cooper et Schindler (2003), sendo consideradas como favoráveis as respostas situadas no nível “CONCORDO” ou superior. A partir deste conceito, são apresentados na tabela 5 os resultados percentuais das respostas emitidas pelos membros da DATAPREV e do INSS consultados, cujas respostas foram consideradas como válidas (127 respondentes) a partir da análise de todos os questionários retornados, constante no APÊNDICE A.

Tabela 5 – Resultado da apuração dos questionários

Questão	CONCORDO OU SUPERIOR	
Q1	91,30%	
Q2	27,60%	
Q3	96,10%	
Q4	97,60%	
Q5	98,40%	
Q6	97,60%	
Q7	73,20%	
Q8	51,20%	
Q9	94,50%	
Q10	94,50%	
Q11	60,60%	
Q12	55,10%	
Q13	100,00%	
Q14	72,40%	
Q15	98,40%	
Q16	77,20%	
Q17	70,90%	
Q18	100,00%	
Q19	100,00%	
Q20	97,60%	
Q21	90,60%	mais de 10 anos
Q22	87,40%	INSS
	12,60%	DATAPREV
Q23	88,20%	INSS
	11,80%	DATAPREV

Fonte: Elaboração própria

No conjunto de questionários analisados, o nível de consenso encontrado foi superior a 50% em todas as questões, exceto na questão 2, que no modelo foi descartada em função de correlação significativa, confirmando a opinião dos respondentes. Uma parcela significativa dos respondentes (90,60%) atua a mais de 10 anos na área de Benefícios por Incapacidade. Do total de respondentes, 87,40% atuam na gestão do INSS, enquanto 12,6% atuam na DATAPREV. Ainda segundo a tabela 5, 88,20% do total de respondentes exercem atividades de Gestão na Administração Central do INSS ou em Agências da Previdência Social (APS), enquanto 11,80% estão envolvidos em atividades de tecnologia da informação do INSS ou da DATAPREV.

3.6 Pesquisa exploratória

A exploração é particularmente útil quando os pesquisadores não têm uma idéia clara dos problemas que serão enfrentados no estudo. Através da exploração podem ser desenvolvidas definições operacionais e conceitos de forma mais clara, além da melhoria do planejamento e o estabelecimento de prioridades (COOPER et SCHINDLER, 2003).

Ainda segundo Cooper et Schindler (2003), a exploração economiza tempo e recursos financeiros, além de auxiliar o pesquisador quando a área de investigação é vaga e o pesquisador necessita fazer uma exploração para saber algo sobre o problema que está sob análise, além de supor que as variáveis podem não ser conhecidas ou não terem sido totalmente definidas.

Para iniciar a pesquisa exploratória e contextualizar de forma quantitativa a distribuição de trabalhadores em atividade remunerada pelas Unidades Territoriais, foi realizada uma consulta ao IBGE ([WWW.ibge.gov.br](http://www.ibge.gov.br)), de onde foi obtida a informação contida na tabela 6 abaixo, onde consta a distribuição da população em função das unidades territoriais.

Tabela 6 - Distribuição da população pelas Unidades Territoriais

Variável = População residente (Pessoas)	
Ano = 2007	
Brasil e Região Geográfica	
Brasil	183.987.291
Norte	14.623.316
Nordeste	51.534.406
Sudeste	77.873.120
Sul	26.733.595
Centro-Oeste	13.222.854

Fonte: IBGE – <http://www.ibge.gov.br/home/geociencias/cartogramas/politico.html>

Na tabela 7 abaixo, segundo o IBGE, consta a distribuição dos cidadãos que exercem atividade remunerada pelas Unidades Territoriais.

Tabela 7 - População ocupada por Unidades Territoriais

Pessoas de 10 anos ou mais de idade, total, economicamente ativas e ocupadas na semana de referência por sexo, situação do domicílio e grupos de idade	
Variável = Pessoas de 10 anos ou mais de idade (Pessoas)	
Sexo = Total	
Situação do domicílio = Total	
Grupos de idade = Total	
Ano = 2000	
Brasil e Região Geográfica	
Brasil	136.910.358
Norte	9.657.961
Nordeste	37.565.737
Sudeste	59.780.090
Sul	20.572.823
Centro-Oeste	9.333.747

Fonte: IBGE em <http://www.sidra.ibge.gov.br/bda/tabela/protabl.asp?z=cd&o=11&i=P>

Para Marques et al. (2006) a pesquisa com objetivo exploratório é adotada na obtenção de maiores informações sobre um determinado tema, inclusive com a finalidade de chegar a problemas específicos e estabelecer hipóteses, visando inclusive servir de referencial teórico para estudos posteriores.

Neste estudo, foi feita uma pesquisa exploratória sobre as informações relativas aos benefícios por Incapacidade constantes nas bases da Previdência Social, buscando quantificar os níveis de significância das variáveis identificadas na enquête realizada junto ao corpo funcional da DATAPREV e ligado às atividades de manutenção e concessão de benefícios por Incapacidade.

Foi desenvolvida uma solução de tecnologia da informação (programa utilizando a linguagem Microsoft Visual Basic versão 6.0) com o propósito de selecionar todos os benefícios por incapacidade constantes nas bases de informações da Previdência Social e transferi-los para o ambiente de tecnologia de baixa plataforma, alvo da condução desta pesquisa, retratando os quantitativos em função da categorização dos benefícios, conforme apresentado na tabela 8.

O objetivo deste estudo é a criação de perfis a partir de benefícios por incapacidade cessados e ativos. Os cessados serão constituídos pelos benefícios informados pela Inspeção do INSS (situação 17 da tabela 7) e pelos benefícios cessados pela Auditoria do INSS (situação 18 da tabela 7), totalizando 63.078 benefícios.

Os benefícios ativos entrarão na composição do modelo e a partir destes foi examinada a correção e o enquadramento destes benefícios. Foi utilizado o conjunto de informações dos benefícios ativos em jan/2008 para compor e validar o modelo proposto, sendo este total de benefícios ativos constituído de 2.373.704 benefícios (situação zero da tabela 8).

Tabela 8 - Posição dos benefícios por incapacidade

Situação	Descrição da Situação	Quantidade	Representatividade
0	ATIVO	2.373.704	16,61%
2	CESSADO	11.375.394	79,62%
3	SUSPENSO	31.153	0,22%
4	SUSPENSO POR MARCA DE ERRO	359	0,00%
5	CESSADO POR CESS DO ORIGEM	25.916	0,18%
6	SUSPENSO P/ SUSP DO ORIGEM	235	0,00%
7	SUSPENSO PELO CONPAG	2.984	0,02%
8	CESSADO PELO SISOBI	408.234	2,86%
11	RECEBENDO MENSALID DE RECUPER 18 MESES	100	0,00%
12	SUSPENSO REVISAO RUR/URB	1.971	0,01%
13	SUSPENSO POR ERRO DE NOME (POVOAMENTO)	3	0,00%
16	SUSPENSO PELA AUDITORIA	472	0,00%
17	CESSADO PELA INSPETORIA	57.716	0,40%
18	CESSADO PELA AUDITORIA	5.362	0,04%
21	SUSPENSO REVISAO RURAL/95	230	0,00%
22	CESSADO REVISAO RURAL/95	659	0,00%
23	SUSPENSO PELO SISOBI	1.712	0,01%
24	CANCELADO PELA AUDITORIA	341	0,00%
Total		14.286.545	100,00%

Fonte: Bases de dados da Previdência Social

A identificação dos benefícios por incapacidade concedidos pela Previdência Social, é feita através de um classificador numérico definido como *ESPÉCIE*. Sua

totalidade está descrita no APÊNDICE B. Os valores das espécies contemplam todos os benefícios gerenciados pelo RGPS. O alvo deste estudo engloba 04 (quatro) espécies da área de benefícios por incapacidade em função de serem as mais significativas em termos do volume de concessões e dos valores pagos pela Previdência Social Brasileira.

Na tabela 9 são apresentados os benefícios por incapacidade, alvo do estudo, distribuídos pelo classificador espécie, utilizado pela Previdência Social.

Tabela 9 - Distribuição dos benefícios por incapacidade por espécie

Espécie	Descrição da Espécie	Quantidade	Representatividade
2	PENSAO POR MORTE ACIDENTARIA-TRAB. RURAL	300	0,00%
4	APOSENTADORIA POR INVALIDEZ-TRAB. RURAL	25.373	0,18%
5	APOSENT. INVALIDEZ ACIDENTARIA-TRAB.RUR.	509	0,00%
6	APOSENT. INVALIDEZ EMPREGADOR RURAL	403	0,00%
10	AUXILIO DOENCA ACIDENTARIO - TRAB. RURAL	89.890	0,63%
13	AUXILIO DOENCA - TRABALHADOR RURAL	59.653	0,42%
25	AUXILIO RECLUSAO	42.185	0,30%
31	AUXILIO DOENCA PREVIDENCIARIO	9.221.814	64,55%
32	APOSENTADORIA INVALIDEZ PREVIDENCIARIA	2.779.115	19,45%
33	APOSENTADORIA INVALIDEZ AERONAUTA	172	0,00%
34	APOSENT. INVAL. EX-COMBATENTE MARITIMO	296	0,00%
36	AUXILIO ACIDENTE PREVIDENCIARIO	9.521	0,07%
50	AUXILIO DOENCA EXTINTO PLANO BASICO	4	0,00%
51	APOSENT. INVALIDEZ EXTINTO PLANO BASICO	1.080	0,01%
83	APOSENTADORIA POR INVALIDEZ EX-SASSE	304	0,00%
91	AUXILIO DOENCA POR ACIDENTE DO TRABALHO	1.539.738	10,78%
92	APOSENT. INVALIDEZ ACIDENTE TRABALHO	112.047	0,78%
93	PENSAO POR MORTE ACIDENTE DO TRABALHO	101.641	0,71%
94	AUXILIO ACIDENTE	238.208	1,67%
95	AUXILIO SUPLEMENTAR ACIDENTE TRABALHO	64.292	0,45%
Total		14.286.545	100,00%

Fonte: Bases de dados da Previdência Social Brasileira

Para o propósito deste estudo, foram analisadas as espécies **31,32,91 e 92**.

A partir da definição e identificação do conjunto de categorias que delimitarão o escopo do estudo, serão apresentadas informações quantitativas sobre a

distribuição dos benefícios na condição de “cessados”, o detalhamento está descrito na tabela 10.

Tabela 10 - Benefícios por incapacidade cessados distribuídos por espécie

Espécie	Descrição da Espécie	Quantidade	Representatividade
04	APOSENTADORIA POR INVALIDEZ-TRAB. RURAL	42	0,07%
25	AUXILIO RECLUSAO	337	0,53%
31	AUXILIO DOENCA PREVIDENCIARIO	40.026	63,45%
32	APOSENTADORIA INVALIDEZ PREVIDENCIARIA	11.442	18,14%
33	APOSENTADORIA INVALIDEZ AERONAUTA	1	0,00%
34	APOSENT. INVAL. EX-COMBATENTE MARITIMO	1	0,00%
83	APOSENTADORIA POR INVALIDEZ EX-SASSE	2	0,00%
91	AUXILIO DOENCA POR ACIDENTE DO TRABALHO	4.873	7,73%
92	APOSENT. INVALIDEZ ACIDENTE TRABALHO	301	0,48%
93	PENSAO POR MORTE ACIDENTE DO TRABALHO	212	0,34%
94	AUXILIO ACIDENTE	730	1,16%
95	AUXILIO SUPLEMENTAR ACIDENTE TRABALHO	5.111	8,10%
Total		63.078	100,00%

Fonte: Bases de dados da Previdência Social

O conjunto de benefícios por incapacidade na condição de “ativos” que foi utilizado no estudo foi obtido das bases corporativas da Previdência Social através de uma solução de tecnologia da informação (*software* integrante da plataforma *mainframe*) especialmente construída para este propósito e processado no ambiente de baixa-plataforma, chegando-se ao conjunto de informações apresentado na tabela 11.

Tabela 11 - Benefícios por incapacidade ativos por espécie

Espécie	Descrição da espécie	Quantidade de benefícios	Representatividade
02	PENSAO POR MORTE ACIDENTARIA-TRAB. RURAL	245	0,01%
04	APOSENTADORIA POR INVALIDEZ-TRAB. RURAL	11.821	0,50%
05	APOSENT. INVALIDEZ ACIDENTARIA-TRAB.RUR.	352	0,01%
06	APOSENT. INVALIDEZ EMPREGADOR RURAL	139	0,01%
10	AUXILIO DOENCA ACIDENTARIO - TRAB. RURAL	1	0,00%
13	AUXILIO DOENCA - TRABALHADOR RURAL	16	0,00%
25	AUXILIO RECLUSAO	7.260	0,31%
31	AUXILIO DOENCA PREVIDENCIARIO	261.840	11,03%
32	APOSENTADORIA INVALIDEZ PREVIDENCIARIA	1.687.288	71,08%
33	APOSENTADORIA INVALIDEZ AERONAUTA	102	0,00%
34	APOSENT. INVAL. EX-COMBATENTE MARITIMO	39	0,00%
36	AUXILIO ACIDENTE PREVIDENCIARIO	8.014	0,34%
50	AUXILIO DOENCA EXTINTO PLANO BASICO	1	0,00%
51	APOSENT. INVALIDEZ EXTINTO PLANO BASICO	197	0,01%
83	APOSENTADORIA POR INVALIDEZ EX-SASSE	108	0,00%
91	AUXILIO DOENCA POR ACIDENTE DO TRABALHO	21.694	0,91%
92	APOSENT. INVALIDEZ ACIDENTE TRABALHO	86.298	3,64%
93	PENSAO POR MORTE ACIDENTE DO TRABALHO	78.283	3,30%
94	AUXILIO ACIDENTE	171.616	7,23%
95	AUXILIO SUPLEMENTAR ACIDENTE TRABALHO	38.390	1,62%
Total		2.373.704	100,00%

Fonte: Bases de dados da Previdência Social

Após a identificação dos benefícios cessados, constantes na tabela 10 e dos ativos constantes na tabela 11, foi feita uma nova consulta aos especialistas da DATAPREV e INSS sobre as informações obtidas para embasar o estudo. Fomos notificados de que poderíamos deparar-mo-nos com ausência parcial de preenchimento das informações sobre a gestão dos benefícios por incapacidade, já que é do conhecimento dos funcionários incumbidos da gestão dos sistemas e negócios previdenciários, a falta de completeza das informações para benefícios solicitados antes do mês de junho de 1994.

Diante destas informações, definimos que as análises seriam realizadas somente em relação aos benefícios por incapacidade nas situações de ativos ou cessados, que tivessem sido requeridos junto a Previdência Social a partir de junho/1994, momento em que passou a ser utilizada uma nova solução de

tecnologia, apoiada em bancos de dados e regras do negócio atualizadas, denominada como SUB (Sistema Único de Benefícios), no âmbito da Previdência Social.

A partir da diretriz fornecida pelos especialistas para evitar os *missing cases* (dados que são desprezados por serem incompletos em relação às variáveis consideradas no estudo) e os *outliers*, que são as observações atípicas, com uma combinação única de características que podem ser diferentes das outras observações, levando-as à divergência do restante da amostra (CORRAR et al., 2007).

Na análise multivariada, especificamente na regressão logística (*logit*), os dados que não contenham valores válidos em todas as suas dimensões devem ser desconsiderados, bem como os que encontrem-se fora dos domínios de validação (HAIR et. al, 2005). Neste estudo os dados considerados fora dos domínios considerados (*outlier*), serão excluídos do estudo, bem como as observações que não contenham todas as informações necessárias ao estudo (*missing cases*).

Após a realização dos ajustes e consideradas somente as espécies que integraram o estudo, o conjunto de informações relativas aos benefícios cessados ficou reduzido, segundo consta na tabela 12.

Tabela 12 - Benefícios por incapacidade cessados após junho/1994

Especie	Descrição da Espécie	Quantidade	Representatividade
31	AUXILIO DOENCA PREVIDENCIARIO	12.456	74,54%
32	APOSENTADORIA INVALIDEZ PREVIDENCIARIA	795	4,76%
91	AUXILIO DOENCA POR ACIDENTE DO TRABALHO	3.437	20,57%
92	APOSENT. INVALIDEZ ACIDENTE TRABALHO	22	0,13%
Total		16.710	100,00%

Fonte: Bases de dados da Previdência Social

A partir dos critérios de completeza das informações definidos pelos especialistas e da redução do conjunto de benefícios ativos somente às espécies que integrariam o estudo, foram selecionadas as informações relativas aos benefícios ativos e iniciados a partir do mês de junho de 1994, época em que a Previdência Social iniciou a utilização da solução de tecnologia da informação denominada como Sistema Único de Benefícios (SUB).

Após a aplicação destes conceitos sobre as informações relativas aos benefícios ativos, foi obtido o conjunto de informações retratado na tabela 13.

Tabela 13 - Benefícios por incapacidade Ativos iniciados após junho/1994

Espécie	Descrição da Espécie	Quantidade	Representatividade
31	AUXILIO DOENCA PREVIDENCIARIO	261.045	18,37%
32	APOSENTADORIA INVALIDEZ PREVIDENCIARIA	1.077.560	75,85%
91	AUXILIO DOENCA POR ACIDENTE DO TRABALHO	21.638	1,52%
92	APOSENT. INVALIDEZ ACIDENTE TRABALHO	60.430	4,25%
Total		1.420.673	100,00%

Fonte: Bases de dados da Previdência Social

De acordo com Hair et al (2005), a construção das amostras para a geração do processo de Regressão Logística (*logit*) necessita de quantidades proporcionais de elementos em ambos os grupos que serão analisados pelo processo *logit*, buscando assim a estimação por verossimilhança, de seus coeficientes e das probabilidades associadas a cada dimensão dos grupos analisados.

Segundo STEVENSON (2001), a amostragem sistemática é semelhante à amostragem aleatória simples, necessitando que seja elaborada uma lista contendo os itens da população. Se os itens da lista não se apresentam em uma ordem determinada, a amostragem sistemática pode resultar em uma amostra realmente aleatória. No caso da pesquisa, foram utilizadas as inscrições dos beneficiários fornecidas pela Previdência Social (Números de benefícios) nas datas de solicitação dos benefícios. Ainda segundo STEVENSON (2001), o processo de amostragem

probabilística sistemática deverá selecionar cada k-ésimo item da lista, onde k é obtido a partir da divisão da população pelo tamanho da amostra que se deseja selecionar, assim :

$$K = N / n$$

Onde:

N é a quantidade de elementos da população

n é o tamanho da amostra desejada

K é o fator que indicará o intervalo para a seleção de cada elemento da amostra a partir da população.

A partir do quantitativo de ocorrências do grupo de benefícios cessados (16.710), foi realizada uma seleção probabilística sistemática sobre os benefícios ativos (1.420.673) com o propósito de obter uma amostra com um quantitativo de benefícios igual ou próximo ao quantitativo de benefícios cessados (16.710). Para gerar a amostra probabilística sistemática foi desenvolvida uma solução de tecnologia da informação (programa na linguagem *Java*) para selecionar de forma sistemática um benefício ativo no universo de 1.420.673 a cada intervalo de 85 benefícios, ou seja:

$$intervalo = \frac{1.420.673}{16.710}$$

A partir deste processo sistemático probabilístico de seleção da amostra, foram obtidos 16.714 benefícios ativos categorizados nas espécies que compõe o objetivo do estudo, de acordo com a tabela 14.

A amostra para execução do processo de Regressão logística binária passou a constituir-se de 16.714 benefícios ativos e 16.710 benefícios cessados, segundo a tabela 15.

Tabela 14 - Benefícios ativos integrantes da amostra sistemática

Espécie	Descrição da espécie	Quantidade	Representatividade
31	AUXILIO DOENCA PREVIDENCIARIO	3.058	18,30%
32	APOSENTADORIA INVALIDEZ PREVIDENCIARIA	12.694	75,95%
91	AUXILIO DOENCA POR ACIDENTE DO TRABALHO	253	1,51%
92	APOSENT. INVALIDEZ ACIDENTE TRABALHO	709	4,24%
Total		16.714	100,00%

Fonte: Elaboração própria

Tabela 15 - Composição da amostra final da pesquisa

Situação	Espécie	Descrição da Espécie	Quantidade	Totais	%
ATIVO	31	AUXILIO DOENCA PREVIDENCIARIO	3.058		18,30%
ATIVO	32	APOSENTADORIA INVALIDEZ PREVIDENCIARIA	12.694		75,95%
ATIVO	91	AUXILIO DOENCA POR ACIDENTE DO TRABALHO	253		1,51%
ATIVO	92	APOSENT. INVALIDEZ ACIDENTE TRABALHO	709	16.714	4,24%
CESSADO	31	AUXILIO DOENCA PREVIDENCIARIO	12.456		74,54%
CESSADO	32	APOSENTADORIA INVALIDEZ PREVIDENCIARIA	795		4,76%
CESSADO	91	AUXILIO DOENCA POR ACIDENTE DO TRABALHO	3.437		20,57%
CESSADO	92	APOSENT. INVALIDEZ ACIDENTE TRABALHO	22	16.710	0,13%
Total			33.424		

Fonte: Bases de dados da Previdência Social

Após algumas reuniões de validação da pesquisa exploratória demonstrada na tabela 15, foram prestadas informações pela Auditoria do INSS de que os dados utilizados como referência para os benefícios cessados pela Auditoria e Inspetoria e que fundamentam este estudo, não deveriam ser utilizados para estabelecer o perfil dos benefícios com suposição de irregularidade e geração do modelo *logit*, uma vez que vários benefícios haviam sido cessados devido à falta de comparecimento do segurado à Agência da Previdência Social, quando solicitado pelo Órgão de Auditoria e Controle da Previdência Social.

Diante destas afirmativas, foram feitas novas extrações de informações e novos processamentos, buscando selecionar os efetivos motivos constatados que levaram à cessação dos benefícios, segundo a tabela 16.

Tabela 16 - Motivos de cessação dos benefícios por incapacidade

MOTIVO	DESCRIÇÃO DO MOTIVO	Quantidade
12	LIMITE MEDICO	12.995
13	OBITO DO TITULAR DO BENEFICIO	952
19	CESS. PA DEVIDO CESS. BENEF. INST.	2
22	PRORROGACAO DE BENEFICIO ANTERIOR	11
28	TRANSFORMACAO PARA OUTRA ESPECIE	1
29	CONCESSAO DE OUTRO BENEFICIO	1
30	CONSTATAÇÃO DE FRAUDE	86
31	CONSTATAÇÃO IRREGULAR./ERRO ADM.	101
32	DECISAO DE CESSACAO POR RECURSO	2
34	VOLTA AO TRABALHO	46
36	ACUMULACAO INDEVIDA DE BENEFICIOS	23
37	SUSP SISOBI MAIS 6 MESES	1
39	NAO ATENDIMENTO CONVOCACAO INSPET.	459
40	CESS.P/ RECUP.TOTAL DENTRO 5 ANOS	25
41	CESS.P/ RECUP. PARCIAL APOS 5 ANOS	77
44	AUX.DOENCA NAO COMPAREC. A PERICIA	312
49	OBITO INFORMADO PELA AUDITORIA	49
52	ERRO ADM. INFORMADO PELA AUDITORIA	496
53	FRAUDE INFORMADA PELA AUDITORIA	505
54	LIMITE MEDICO INFORMADO P/ PERICIA	3
69	ALTA MEDICA	76
70	RETORNO VOLUNTARIO AO TRABALHO	20
71	ERRO TECNICO	15
72	NAO COMPARECIMENTO	38
77	IRREGULARIDADE DETECTADA PELA AUDITORIA	408

Fonte: Elaboração própria

A partir das informações da tabela 16 e após a realização de algumas reuniões com os membros da Auditoria do INSS e especialistas da área de Benefícios por Incapacidade, chegou-se ao consenso de que os motivos identificados na tabela 16 como os descritores 30, 44, 53 e 77, deveriam retratar com fidedignidade o universo dos benefícios que foram cessados sob alegação de fraude ou irregularidade. O quantitativo de benefícios cessados definido através deste critério passou a ser de 1.311 (hum mil trezentos e onze) benefícios por incapacidade.

Diante da alteração da quantidade de benefícios por incapacidade cessados por irregularidade, um novo processo de amostragem probabilística sistemática foi realizado, com o objetivo de obter uma quantidade similar (1311) de benefícios por

incapacidade na situação de ativos e iniciados após o mês de junho de 1994, como já citado anteriormente, época em que uma nova solução de tecnologia da informação fora adotada, destinada à gestão de Benefícios pela Previdência Social com maior acurácia e controle.

Através da construção e utilização de uma nova solução de tecnologia (programa na linguagem Java) e segundo STEVENSON (2001), o novo processo de seleção de amostragem probabilística sistemática foi desenvolvido e resultou em:

$$K = N / n$$

Onde:

N é a quantidade de elementos da população

n é o tamanho da amostra desejada

K é o fator que indicará o intervalo para a seleção de cada elemento da amostra a partir da população.

Onde:

$$N = 1.420.673$$

$$n = 1.311$$

logo $K = 1.420.673 / 1311$, resultando em um valor de $K = 1.084$ registros.

Como resultado deste processo foram obtidos 1311 benefícios constantes nas bases de dados da Previdência Social na condição de ativos. Após selecionados, foram reunidos aos outros 1311 benefícios na situação de cessados, constituindo a nova amostra sobre a qual foi executado o processo de regressão logística (*logit*), contendo 2.622 benefícios, conduzindo a uma nova pesquisa exploratória.

A tabela 17 apresenta a distribuição dos benefícios por espécie, considerando os critérios informados pela Auditoria do INSS.

Tabela 17 - Distribuição dos benefícios por espécie

Situação	Descrição da Situação	Código da Espécie	Descrição da Espécie	Quantidade	Representa:
0	Ativo	31	AUXILIO DOENCA PREVIDENCIARIO	246	18,76%
0	Ativo	32	APOSENTADORIA INVALIDEZ PREVIDENCIARIA	991	75,59%
0	Ativo	91	AUXILIO DOENCA POR ACIDENTE DO TRABALHO	22	1,68%
0	Ativo	92	APOSENT. INVALIDEZ ACIDENTE TRABALHO	52	3,97%
1	Cessado	31	AUXILIO DOENCA PREVIDENCIARIO	1083	82,61%
1	Cessado	32	APOSENTADORIA INVALIDEZ PREVIDENCIARIA	154	11,75%
1	Cessado	91	AUXILIO DOENCA POR ACIDENTE DO TRABALHO	72	5,49%
1	Cessado	92	APOSENT. INVALIDEZ ACIDENTE TRABALHO	2	0,15%
Total				2622	

Fonte: Elaboração própria

Na tabela 18 é apresentada a distribuição da amostra pelo sexo dos segurados.

Tabela 18 - Distribuição dos segurados por sexo informado

Situação	Descrição da Situação	Código do sexo	Descrição do sexo	Quantidade	Representa
0	Ativo	1	MASCULINO	778	59,34%
0	Ativo	3	FEMININO	533	40,66%
1	Cessado	1	MASCULINO	917	69,95%
1	Cessado	3	FEMININO	394	30,05%
Total				2622	

Fonte: Elaboração própria

Para Cooper et Schindler (2003), a pesquisa exploratória termina quando os pesquisadores já conseguiram estabelecer as principais dimensões da tarefa da pesquisa, tais como:

- Definir um conjunto de questões investigativas que poderão ser utilizadas para guiar o planejamento detalhado da pesquisa;
- Descobrir que determinadas hipóteses são possibilidades tão remotas que podem ser ignoradas em qualquer estudo futuro;
- Concluir que não é necessário ou não é viável fazer outras pesquisas adicionais.

Para DESLANDES et al. (1998:31), a fase exploratória de uma pesquisa é um dos momentos mais importantes, podendo ser considerada como uma pesquisa exploratória (MINAYO,1992), compreendendo várias fases da trajetória da investigação, tais como: a escolha do tópico da investigação; a delimitação do problema; a definição do objeto e dos objetivos; a construção de um marco teórico conceitual; a escolha dos instrumentos de coleta de dados e a exploração de campo.

De acordo com MINAYO et al. (1998):

“Formalmente a fase exploratória de uma pesquisa termina quando o pesquisador definiu o seu objeto de pesquisa, construiu o marco teórico conceitual a ser empregado, definiu os instrumentos de coleta de dados, escolheu o espaço e o grupo de pesquisa, definiu a amostragem e estabeleceu estratégias para entrada no campo.”

Ainda segundo MINAYO et al. (1998), a fase exploratória baseia-se em muitos esforços, em uma pesquisa bibliográfica disciplinada e crítica, através da elaboração de fichamentos e resumos e do estabelecimento de um diálogo reflexivo entre a teoria e o objeto de investigação selecionado pela pesquisa. Nesta fase também deve ser verificado o conhecimento sobre o assunto, existente e registrado em meios aceitos sob a ótica epistemológica, deve ainda ser embasada por conceitos de humildade, através do reconhecimento de que todo conhecimento científico tem sempre um caráter. Deve ser aplicado o conceito de proximidade, que se baseia no fato de que os conceitos utilizados no estudo são fundamentados em outros conhecimentos, sobre os quais serão realizados questionamentos, aprofundamentos ou críticas. Deve-se considerar também, que o conhecimento sobre a totalidade do objeto pesquisado é inacessível e que as idéias ou explicações que fazemos da realidade estudada são sempre imprecisas, e parciais em relação à própria realidade (MINAYO et al.,1998).

Ainda segundo MINAYO et al. (1998), o objeto pesquisado deve estar vinculado à vida real, e “... um problema intelectual surge a partir de sua existência na vida real e não “espontaneamente”.”.

Neste estudo, foi identificado que o conjunto de informações obtidas e os dados observados e analisados, são evidências e fundamentos suficientes para prosseguir neste estudo de caso, dando continuidade à pesquisa vislumbrando o teste das hipóteses através da abordagem quantitativa.

Para obter a opinião dos respondentes deste estudo, composta pelos técnicos ligados à concessão e manutenção dos benefícios por Incapacidade na Previdência Social, sendo estes, funcionários da DATAPREV ou da MPAS, foram enviados 200 (duzentos) questionários, segundo o APÊNDICE A, por e-mail com comunicação prévia da Diretoria de Relacionamento, Desenvolvimento e Informações da DATAPREV e Órgãos de Gestão do INSS, definindo o propósito, embora os questionários tenham sido enviados juntamente com uma carta explicando a utilização das informações prestadas, medida justificada por Lakatos e Marconi (2007a, p. 203), além de ser esperado um baixo retorno das respostas, também segundo Lakatos e Marconi (2007a), quando alegam que em média os questionários expedidos pelo pesquisador alcançam 25% de devolução, e que:

Junto com o questionário deve-se enviar uma nota ou carta explicando a natureza da pesquisa, sua importância e a necessidade de obter respostas, tentando despertar o interesse do recebedor, no sentido de que ele preencha e devolva o questionário dentro de um prazo razoável.

Segundo Cooper et Schindler (2003), a construção das escalas de medida pode ser organizada através de cinco métodos, a saber:

1. Arbitrário

A escala é feita sob medida para mensurar uma propriedade ou um indicador;

2. Consenso

Respondentes (juízes) avaliam e opinam sobre os itens que serão incluídos;

3. Análise de item

As escalas de mensuração são testadas com uma amostra de respondentes;

4. Fatoração

As escalas são construídas a partir de inter-correlações de itens de outros estudos;

5. Cumulativa

A partir da conformidade com uma classificação de itens com poder discriminatório ascendente ou descendente;

Na condução deste estudo foi adotada a escala de consenso, através da qual os especialistas da DATAPREV e INSS, foram os juízes e definiram quais as variáveis que seriam analisadas pelo modelo proposto (COOPER e SCHINDLER, 2003).

O processo de autorização para a utilização das informações neste trabalho foi precedido de várias justificativas e autorizações dos integrantes de cargos de gestão, tanto do INSS e DATAPREV (Diretor de Relacionamento, Coordenação Geral da Auditoria Interna) quanto de integrantes do MPAS (Coordenador Geral de Modernização dos Sistemas Previdenciários), a partir de uma justificativa formalizada definindo os propósitos, objetivos e o nível de relevância para a Previdência Social e sociedade em geral.

Tais controles e exigências justificam-se pelo sigilo e confidencialidade que tais informações de cidadãos necessitam, trata-se de informações governamentais e estratégicas, razões pelas quais informações que possibilitem a associação ou a identificação de cidadãos, serão omitidas deste estudo e de sua publicação.

O conjunto de informações utilizado neste estudo refere-se às bases de dados que contêm as informações dos Benefícios, informações apresentadas no processo de concessão (Sócio-econômicas e laborativas) e Informações dos Beneficiários Previdenciários (Sociais).

As informações são mantidas por sistemas que utilizam a tecnologia *Mainframe* (Fornecedor Unisys). Durante o planejamento da obtenção das informações, foi definido que o processo ocorreria através da transferência de informações para o ambiente de Micro-Informática, devido aos significativos volumes de informações existentes (históricas e de manutenção dos benefícios).

Foi feito um estudo para verificar as transformações necessárias para comportar as análises estatísticas (Regressões Logísticas e Estatísticas Descritivas) sobre as dimensões alvo. A solução adotada deveria ser comportada por equipamentos de baixo poder de processamento, garantindo a possibilidade de continuação do estudo por outros pesquisadores ou interessados no tema, sem a necessidade de estrutura de Tecnologia da Informação específica ou com requisitos de processamento robustos ou de alto custo.

A partir de um estudo exploratório nas bases de informações da Previdência Social, foram verificados os volumes de informações disponíveis, a saber: Benefícios – 65.000.000, Concessão – 45.000.000 e Beneficiários – 54.000.000, segundo o quadro 3.

Quadro 3 - Volume de informações disponível

Tipos de informações	Volume disponível
Benefícios	65.000.000
Concessão	45.000.000
Beneficiários	54.000.000
Volume total	164.000.000

Fonte: Fonte: Elaboração própria

O foco inicial foram as informações relativas aos Benefícios concedidos na categoria de Benefícios por Incapacidade, apresentados no quadro 4.

Quadro 4 - Distribuição dos Benefícios por Incapacidade

Acidentários	Quantidade	Valor
Aposentadorias por Invalidez	147.457	98.195.116
Pensões por Morte	128.985	76.887.312
Auxílio-Doença	137.830	110.371.098
Auxílio-Acidente	269.724	92.936.563
Auxílio-Suplementar	83.940	9.203.181
Total	767.936	387.593.269

Fonte: Anuário Estatístico da Previdência Social

O processo de extração das informações constantes na base de benefícios, ocorreu através de uma aplicação de TI (Tecnologia da Informação), que executou o processamento no ambiente *Mainframe* e transformou as informações constantes nas bases de dados para um formato possível de ser transferido para o ambiente de Micro-informática.

Foram selecionadas no ambiente Mainframe todas as informações relativas às categorias de Benefícios por Incapacidade. Foi realizada a extração das informações de 500.000 (quinhentos mil) benefícios para contemplar o estudo exploratório.

O gerenciador de bancos de dados Mysql AB (<http://www.mysql.com>) foi selecionado devido ao seu desempenho (*performance*), exigência de poucos recursos do equipamento e baixa alocação de espaço em disco pelos arquivos contendo as informações, além desta tecnologia ter a sua utilização e licenças disponíveis ao público em geral (*GPL – General Public Licence*) sem custos.

Foi realizada a transferência das informações sobre 14.286.545 benefícios relativos aos Benefícios por Incapacidade, através de uma aplicação de TI especialmente desenvolvida para este propósito no ambiente *Mainframe* (Unisys).

O tratamento das informações na plataforma de micro-informática (verificações de completeza e separação em arquivos menores) e armazenamento em base de dados Mysql foi realizada em um micro-computador (Notebook, Processador AMD Turion 1.6ghz, 2gb de memória e 100gb de disco rígido (*Hard Disk*)), dotado também de um meio de armazenamento externo de 160gb.

Foi aplicado o processo Hosmer-Lemeshow (60% testes x 40% validação) sobre a totalidade das informações dos benefícios constantes nas bases de dados da Previdência Social e que figuravam na situação de “Cessados pela Auditoria ou Inspeção do INSS”, devido à constatação de irregularidades pelas Auditorias do MPAS/INSS e DATAPREV e Inspeção do MPAS/INSS.

Em uma fase seguinte, este conjunto de dados foi dividido em duas sub-amostras sobre as quais foram processadas regressões logísticas, uma com o propósito de gerar o modelo (60%) (coeficientes e interceptos) e outra com o propósito de validar o modelo obtido na primeira, sendo esperado do método que este comportamento seja refletido na amostra de validação (40%) (HAIR et al., 2005).

3.7 Regressão Logística

Segundo Hair et al. (2005, p.231), quando a variável dependente tem apenas dois grupos, a regressão logística pode ser preferida por diversas razões, este estudo busca a identificação de benefícios que possam ou não apresentar irregularidades, portanto, deverão ou não integrar os processos de Auditoria. Ainda segundo Hair e al. (2005, p. 232), a regressão logística se aproxima da tarefa de predição de probabilidade de maneira mais semelhante à encontrada em regressão

múltipla, diferindo no sentido de que a Regressão Logística prevê diretamente a probabilidade de um evento ocorrer.

Em relação ao domínio dos valores resultantes de um processo de Regressão Logística, Hair e al. (2005, p. 232), afirmam que os valores de probabilidades podem ser quaisquer valores entre zero e um, mas o valor previsto deve ser limitado, de modo a recair no intervalo de zero a um, sendo a relação assumida entre as Variáveis Independentes (VI) e Variável Dependente (VD), lembrando uma curva em forma de “S”, definida como *relação logit*.

A regressão múltipla emprega o método dos mínimos quadrados, o qual minimiza a soma de diferenças quadradas entre os valores reais e os valores previstos para a variável dependente e que a natureza não-linear da transformação logística demanda que um outro procedimento, o de máxima verossimilhança, seja usado de forma iterativa para encontrar estimativas “mais prováveis” para os coeficientes, através do ajuste da curva logística aos dados reais (HAIR et al. ,2005, p. 232) .

Em relação à condução da técnica de Regressão logística (*logit*) e como descrito em Hair e al. (2005, p. 234) “Um modelo nulo, o qual é semelhante a calcular a soma total de quadrados usando a média, fornece o ponto de referência para a comparação”, detalhando ainda que “... há outros métodos para avaliar o caráter não-métrico da Variável Dependente (VD), tais como Matrizes de Classificação e o processo de Hosmer & Lemeshow”.

Para Corrar et .al. (2007), um ponto em relação ao qual a literatura ainda não apresenta consenso é a quantidade de observações necessárias à realização de inferências de boa qualidade. Os autores afirmam que uma regra razoável é obter

um número de observações trinta vezes a quantidade de parâmetros que se deseja estimar.

Uma das principais vantagens associadas ao uso da Regressão Logística está na relativa facilidade com que ela pode explicar e prever a ocorrência de determinados fenômenos nas diversas áreas do conhecimento, tais como economia, administração, contabilidade, sociologia e medicina. Pode-se afirmar que o modelo logístico se presta a elucidar dois grandes objetivos: identificar a que grupo certos objetos, pessoas ou fenômenos pertencem e estimar a probabilidade de que eles possam se enquadrar nesta ou naquela categoria (CORRAR, 2007).

Um modelo de regressão pode ser definido como uma equação matemática em que se expressa o relacionamento de variáveis. Nestes modelos, define-se uma variável dependente (Y), ou variável de saída, e procura-se verificar a influência de uma ou mais variáveis ditas variáveis independentes, causais ou explicativas (X's) sobre esta variável dependente (HAIR et al.,2005; CORRAR et al., 2007; TABACHNICK et FIDELL,2001).

Para Corrar et al. (2007, p. 287) e Hair et. al (2003), no modelo logístico o coeficiente de cada variável independente exerce efeitos sobre a quantidade de *logit* (logaritmo natural de chance) e mede a mudança que ocorrerá na variável dependente por unidade de variação ocorrida na variável independente, como demonstrado a seguir.

$$\ln\left(\frac{P}{1-p}\right) = b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k$$

Onde:

ln: logaritmo natural

p: Probabilidade de ocorrência do fenômeno

1-p: Probabilidade de não-ocorrência do fenômeno

b_i: são os coeficientes de regressão;

x_i : são as variáveis independentes;

Segundo Corrar et al. (2007), a equação abaixo permite o cálculo da probabilidade relativa à ocorrência de determinado evento, sendo necessário para tanto, o cálculo dos coeficientes e as variáveis preditoras independentes que integram o modelo final, a partir dos critérios de redução do *likelihood value*, testes de significância do modelo e da significância dos seus coeficientes.

$$P = \frac{I}{1 + e^{-(b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k)}}$$

Para Tabachnick e Fidell (2001), a regressão logística consiste em um tipo de regressão aplicável e preferida quando se tem uma variável dependente categórica dicotômica, ou seja, uma variável nominal ou não métrica que possui apenas dois grupos ou classificações como resultados possíveis como, por exemplo, alto ou baixo, homem e mulher, sim ou não etc.

A partir da obtenção dos coeficientes e da equação logística, o mesmo modelo foi aplicado aos 40% dos dados restantes e selecionados previamente, buscando a confirmação do modelo criado. Caso estes coeficientes tenham significância estatística, segundo Hair e al. (2005, p. 235) o modelo poderá ou não ser ajustado, ainda segundo Hair e al. (2005, p. 235) “... *testar a hipótese dos coeficientes serem diferentes de zero...*”, uma vez que coeficientes iguais a zero significam que a razão de desigualdade não muda e a probabilidade não é afetada.

O teste de significância dos coeficientes foi feito através da Estatística Wald, que fornecerá a significância estatística de cada coeficiente, sendo o coeficiente Wald obtido a partir de $W = (\beta_i / SE_{\beta_i})^2$, onde β_i corresponde ao coeficiente e SE corresponde ao Erro padrão identificado para a variável, com um grau de liberdade (SINICICH, 1996).

A utilização de Regressão Logística foi fundamentada no fato de que algumas dimensões que compõem os benefícios por incapacidade são de natureza categórica e segundo Hair e al. (2005, p. 232), “... *os modelos lineares de regressão não podem acomodar tal relação, uma vez que ela é inerentemente não-linear* “ e que “... *o termo de erro de uma variável discreta segue a distribuição binominal ao invés da distribuição normal, invalidando assim todos os testes estatísticos baseados no pressuposto de normalidade*“. Ainda segundo Hair e al. (2005, p. 232), “... a variância de uma variável dicotômica não é constante, o que também cria casos de heteroscedasticidade”, conforme afirmado também por Tabachnick e Fidell (2001).

Na equação de regressão logística, para verificar-se o efeito ou poder de discriminação de cada uma das variáveis independentes com relação à variável dependente, são calculados os coeficientes de regressão. O cálculo dos coeficientes do modelo é feito através da maximização da função de verossimilhança que calcula a probabilidade de que um evento ocorra (MENARD, 1995; TABACHNICK et FIDELL, 2001; HAIR et al., 2005). Este procedimento é equivalente a minimizar a função logaritmo de verossimilhança (-2LL). Deve ser verificado o poder de ajuste da equação, ou seja, verificar o quanto as variáveis independentes explicam a variável dependente, a medição do poder de influência sobre a variável dependente (TABACHNICK et FIDELL, 2001; HAIR et al., 2005).

Um modelo com bom ajuste terá um valor baixo para -2LL, sendo que o valor mínimo é 0 (zero). Um modelo com ajuste perfeito terá como resposta um valor de verossimilhança igual a 1 (um) e, portanto, -2LL será igual a 0 (zero) (TABACHNICK et FIDELL, 2001; HAIR et al., 2005).

O valor da verossimilhança também pode ser comparado entre equações, onde a diferença representa a mudança no ajuste preditivo de uma equação para

outra. O teste qui-quadrado para a redução no valor do logaritmo da verossimilhança fornece uma medida de melhoria devido à introdução da(s) variável(eis) independente(s) (TABACHNICK et FIDELL, 2001; HAIR et al.,2005; CORRAR et al., 2007).

No estudo em questão, foi definido um modelo logístico, baseado no comportamento das variáveis selecionadas, buscando identificar a probabilidade de um Benefício por Incapacidade mantido pela Previdência Social, estar correto ou incorreto. A variável dependente (Y) indica se o benefício está “Correto” (= 0) ou “Incorreto” (= 1) e a série de indicadores (X_1, \dots, X_p) constitui o conjunto de variáveis independentes a serem analisadas.

Uma das vantagens consideráveis da regressão logística é poder identificar apenas se um evento ocorreu para então usar um valor dicotômico da variável dependente. A partir deste valor dicotômico, o procedimento prevê sua estimativa de probabilidade de ocorrência. Se a probabilidade prevista for maior que 0,50, então a previsão será “sim”, caso contrário será “não”. A regressão deriva seu nome da transformação logística usada com a variável dependente (HAIR et al., 2005;CORRAR et al.,2007;SINCICH,1996; TABACHNICK et FIDELL, 2001).

O procedimento calcula o coeficiente logístico e compara a probabilidade de um evento ocorrer com a probabilidade de não ocorrer (razão de chance). Essa razão de desigualdade pode ser expressa na figura 3.

$$\frac{\text{Prob}(\text{evento ocorrer})}{\text{Prob}(\text{evento não ocorrer})} = e^{\beta_0 + \beta_1 X_{1i} + \dots + \beta_n X_{ni}}$$

Figura 3 - Razão de chance da Regressão Logística

Os coeficientes estimados ($\beta_0, \beta_1, \beta_2, \dots, \beta_n$) são as medidas das variações na proporção das probabilidades, denominada de razão de desigualdade. Um coeficiente positivo indica aumento da probabilidade, ao passo que um valor negativo representa diminuição da probabilidade prevista, além da influência ser medida através do antilogaritmo do coeficiente β_i (e^{β_i}) de cada variável (SINCICH,1996).

Para Sincich (1996), o valor de e^{β_i} representa a mudança nas chances de ser observada uma alteração nas chances de observação do evento previsto pelo modelo (variável dicotômica binária = 1), quando a variável independente associada a este coeficiente é elevada em uma unidade, mantidas as demais variáveis constantes.

O teste Hosmer e Lemeshow ou teste HL tem a finalidade de avaliar a validade preditiva do modelo de Regressão Logística. É baseado não no valor de verossimilhança, mas na visão real da variável dependente (HOSMER et LEMESHOW,1989).

Segundo Hair et al. (2005, p.234), para a execução do teste Hosmer e Lemeshow a amostra é dividida em 10 classes iguais, com no mínimo 5 observações e nunca menos de uma, o número de eventos reais e previstos são comparados em cada classe através da estatística qui-quadrado. O resultado do teste é de ampla precisão e não baseia-se somente no valor da verossimilhança obtida, mas sim na real comparação entre os grupos de cada classe, garantindo a previsão sobre a representação dos casos observados pelo modelo.

Considerando-se Y como o valor real da variável e Y_p como o valor previsto, o teste HL é feito com o intuito de medir a proximidade de ambos. A hipótese nula (hipótese de teste) indica que não há diferença significativa entre o valor real e o

valor previsto, ou seja, equivale a dizer que o modelo tem bom poder de ajuste (HAIR et al., 2005).

Quanto menor o valor da diferença entre Y e Y_p , mais os valores previstos se aproximam dos reais e, portanto, melhor desempenho preditivo tem o modelo. Desta forma, um fator positivo a favor do modelo é quando se aceita a seguinte hipótese nula: $H_0: Y - Y_p = 0$ (CORRAR, 2007).

Em relação aos testes de significância do modelo, indicam que o teste *qui-quadrado* (*chi-square*) deve ser utilizado, uma vez que neste estudo, diversos dados nominais estão compondo as dimensões dos benefícios, além de “diferenças entre distribuição observada dos dados e a distribuição esperada com base na hipótese nula”, bem como pelo fato dos coeficientes serem calculados sobre valores reais e não com percentagens (COOPER e SCHINDLER, 2003, p. 397; TABACHNICK et FIDELL, 2001; HAIR et al., 2005; CORRAR et al., 2007).

A partir da obtenção do valor do logaritmo de verossimilhança (-2LL), foi utilizado o procedimento *stepwise* para a inclusão parcial de variáveis ao modelo, utilizando critérios de entrada das variáveis. Segundo Hair et al. (2005, p. 261); Tabachnick et Fidell (2001), “diversos critérios podem ser empregados para orientar a entrada: maior redução no valor -2LL, maior coeficiente de Wald ou maior probabilidade condicional.”.

Para Corrar et al. (2007), o modelo *logit* gerado deverá ser testado em sua completeza e significância a partir do exame dos indicadores pseudo- r^2 , tais como o McFadden's- r^2 , o Cox-Snell- r^2 e o Nagelkerke- r^2 . Estes indicadores servem para comparar o desempenho de modelos concorrentes, sendo adotado o modelo que indicar os maiores valores dos indicadores pseudo- r^2 , por expressarem a maior

variação de logit em relação à variável dependente em função dos coeficientes e das variáveis explicativas (independentes) incluídas no modelo (HAIR et al., 2005).

Segundo Corrar et al. (2007) e Hair et al. (2005) o coeficiente pseudo- r^2 de McFadden é obtido a partir comparação entre o *likelihood value* do modelo inicial (só com a constante) e o modelo seguinte (após a inclusão das variáveis explicativas, sendo representado por:

$$\text{McFadden's-}r^2 = \frac{-2LL_{\text{nulo}} - (-2LL_{\text{modelo}})}{-2LL_{\text{nulo}}}$$

Onde:

McFadden's- r^2 : Redução de -2LL entre o modelo inicial e o atual

-2LL_{nulo} : Log de verossimilhança só com a constante

-2LL_{modelo} : Log de verossimilhança após a inclusão das variáveis explicativas

A medida pseudo- R^2 de Cox e Snell possibilita a avaliação da melhoria na comparação entre modelos, devendo ser considerado como dotado de maior capacidade preditiva àquele modelo que apresentar o maior valor desta medida, que considera a redução do valor do log de verossimilhança através de uma estatística qui-quadrado, tendo o seu *range* compreendido entre 0 e quase 1, não atingindo o limite superior, segundo Minussi (2001) apud PEREIRA et NESS-JUNIOR (2008) o valor máximo da medida Cox e Snell é 0,75 (HAIR et al., 2005; CORRAR et al., 2007; COX et SNELL, 1989).

A medida Nagelkerke, destina-se à mensuração da capacidade preditiva geral do modelo logístico através da comparação entre modelos, sendo considerado o melhor modelo aquele que apresente o maior valor desta medida, cujo *range* situa-se entre 0 e 1 (inclusive) (HAIR et al., 2005; CORRAR et al., 2007; NAGELKERKE, 1991).

Para garantir a significância estatística deve ser feito um teste qui-quadrado para verificar a variação no valor -2LL, além disto, deve ser utilizada a medida

Hosmer e Lemeshow de ajuste geral, buscando diferenças estatisticamente significantes entre as classificações observadas e previstas (HAIR et al., 2005, p. 264; TABACHNICK et FIDELL, 2001).

Estas duas medidas, em combinação, fornecem suporte para a aceitação dos modelos de duas variáveis. Os coeficientes estimados para as variáveis independentes (β_1) e a constante (β_0) podem ser avaliados em termos de significância estatística através da estatística Wald, exceto nos casos em que os coeficientes sejam extremamente grandes, situação em que o escore estatístico é empregado (CORRAR, 2007).

Os coeficientes são estatisticamente significantes no nível 0,01, apesar da constante ser significativa apenas no nível 0,10. Nestas condições as variáveis individuais são significantes e devem ser interpretadas (CORRAR, 2007).

Segundo Hair et al (2005, p. 264), a medida final do ajuste de um modelo é o valor de Hosmer e Lemeshow, que mede a correspondência entre os valores reais e os previstos da variável dependente (VD), sendo também aceito como um bom ajuste o fato de possuir um valor qui-quadrado não-significante, conforme apresentado na figura 4.

$$\chi^2_{square} = \frac{\sum (\text{observado} - \text{esperado})^2}{\text{esperado}}$$

Figura 4 - Demonstração do modelo qui-quadrado

Para Corrar (2007), o êxito alcançado pela técnica de Regressão Logística na literatura especializada está baseado em certos fatores, tais como:

- Acolher com mais facilidade as variáveis categóricas;

- Mostra-se mais adequada à solução de problemas que envolvam estimação de probabilidades, pois trabalha com uma escala de resultados entre 0 (zero) e 1 (um);
- Requer um reduzido número de suposições iniciais em relação a outras técnicas;
- Admite simultaneamente variáveis independentes (VI) métricas e não-métricas;
- Facilita a construção de modelos destinados à previsão de riscos em diversas áreas do conhecimento (*score*);
- Os resultados do modelo podem ser interpretados com relativa facilidade, pois a lógica do modelo assemelha-se às outras técnicas conhecidas, tais como a Regressão Linear e a facilidade na utilização de ferramentas de TI, já que vários pacotes estatísticos possuem estas funcionalidades embutidas.

As organizações têm utilizado a Regressão Logística em larga escala em função de julgarem esta técnica muito adequada para: explicar riscos de crédito, explicar tendências, prever riscos de falência e outros semelhantes resultados da análise de dados. O baixo grau de complexidade operacional e a ausência de restrições rígidas, justificam o avanço da utilização desta técnica, em relação às técnicas similares, tais como a Análise de Discriminante, que pressupõe uma distribuição normal para as variáveis independentes (CORRAR, 2007).

4 APLICAÇÃO PRÁTICA

4.1 Apresentação do modelo gerado

O modelo resultante da pesquisa foi aplicado sobre o conjunto de benefícios por Incapacidade constantes na situação de “ativos”, referentes aos segurados que recebem valores relativos à inatividade, devido à impossibilidade de exercer atividades remuneradas durante o período de validade do benefício.

Foi adotada a técnica de levantamento de dados como método de coleta, tendo sido realizada através da análise dos trabalhos já executados pela Auditoria do MPAS e DATAPREV e da comparação com os resultados obtidos sobre as informações extraídas do acervo da Previdência Social, após a aplicação do método estatístico proposto por este estudo.

Após a verificação do nível de precisão atingido pelo modelo, este foi extrapolado para todo o acervo de benefícios das categorias de *benefícios por incapacidade* da Previdência Social Brasileira, buscando identificar os benefícios que encontram-se em situação divergente, considerando os fatores comuns identificados pelo modelo e as informações relativas às outras categorias de benefícios onde o modelo foi aplicado, uma vez que o modelo em estudo, por ser um estudo de caso, destina-se apenas ao estabelecimento dos scores sobre as quatro espécies que compõem o universo do estudo.

A partir desta posição, os dados já constantes do ambiente de Computação Pessoal (PC), foi utilizado um *software* estatístico para executar a técnica inferencial de Regressão Logística, utilizando acessos via *ODBC (Open Database Connectivity)* à base de dados MySql AB, onde estavam armazenadas as informações relativas aos benefícios por incapacidade.

Foram selecionados todos os benefícios nas situações de “CESSADO” pela Auditoria do INSS e pela Inspeção do INSS, com o propósito de confirmar a situação dos mesmos a partir da análise das características (dimensões) que integraram os benefícios (variáveis explicativas), tendo como resultado uma variável **dicotômica** (INCORRETO ou CORRETO), segundo afirmado por Hair e al. (2005, p. 235), *“o pesquisador que se defronta com uma variável dicotômica não precisa apelar para métodos elaborados para acomodar as limitações da Regressão Múltipla e nem precisa empregar Análise Discriminante”*.

Para realizar este processo, o conteúdo da amostra foi dividido em 10 classes iguais com um mínimo de 20 observações em cada classe, segundo Hosmer e Lemeshow apud Hair et al. (2005). Em seguida, foram selecionadas as dimensões constituintes dos benefícios por Incapacidade (60% dos casos, ainda segundo Hosmer e Lemeshow) e sobre estas informações foi realizado o processo de Regressão Logística (*logit*).

Uma vez obtido e validado, o modelo foi aplicado sobre todos os benefícios por incapacidade, requisitados à Previdência Social após junho/1994.

A partir da significância do modelo, este poderá integrar o Sistema de Concessão e Manutenção de Benefícios da Previdência Social, buscando satisfazer a função de Monitoramento Contínuo (detecção de irregularidades no momento de suas ocorrências e relato imediato aos órgãos de controle), além de possibilitar a seleção de amostras para análises mais profundas de conformidade dos benefícios, partindo dos níveis de risco em relação à conformidade, obtidos através da aplicação prévia do modelo, que tem como objetivo gerar *scores* de risco para todos os benefícios por incapacidade, já concedidos ou em processos de concessão.

Esta definição contribuirá para a extinção de amostras não-probabilísticas nos exames de conformidade pelos Órgãos de Controle, garantindo ações efetivas e direcionadas, que poderão conduzir à melhoria da Gestão das Auditorias, Inspetorias e outros Órgãos de Controle Governamental sobre a gestão dos benefícios por Incapacidade no âmbito da Previdência Social Brasileira.

Após o processamento dos questionários, constantes no APÊNDICE C, tornou-se possível a realização da seleção das características dos benefícios, identificadas através de um *survey de experiência* com os especialistas do MPAS e DATAPREV, a partir de informações sobre segurados que encontram-se como titulares de benefícios na categoria de Benefícios Acidentários.

As informações foram obtidas através de um processo de extração sobre o conjunto de 14.286.545 benefícios contidos no cadastro de benefícios em Janeiro/2008, tendo sido selecionados os benefícios cessados por irregularidades constatadas pelos Órgãos de Inspetoria e Auditoria do INSS (Instituto Nacional do Seguro Social).

Para este estudo, foram selecionados os 1.211 benefícios cessados e outros 1.211 benefícios na situação de Ativos a partir do conjunto de 1.420.673 benefícios por incapacidade ativos concedidos e mantidos pelo INSS e concedidos após junho de 1994.

Para executar um estudo de Regressão Logística, deve ser verificada a correlação entre as variáveis métricas explicativas (Variáveis Independentes), característica denominada como multicolinearidade, evitando que o modelo seja influenciado pelas correlações e os processos *stepwise* sejam incorretamente aplicados na retirada ou na inclusão de variáveis em modelos com alto poder preditivo (HAIR et al (2005); TABACHNICK e FIDELL (2001); CORRAR (2007).

O quadro 5, apresenta as multicolinearidades identificadas quando foi executado o exame das correlações existentes entre as variáveis independentes métricas a partir do software estatístico (opção *analyse, correlate, bivariate*).

Quadro 5 - Análise de multicolinearidade da variáveis independentes

		D_SEG_DIB	ID_SEG_2008	TP_Contrib	QT_DEP_IR	QT_SM_RMI	VL_MR_ATU	QT_REVISAO_RMI	Meses_Receb
ID_SEG_DIB	Pearson Correlation	1	,941(**)	,179(**)	-,006	-,162(**)	-,034	,000	,194(**)
	Sig. (2-tailed)		,000	,000	,772	,000	,081	,986	,000
ID_SEG_2008	Pearson Correlation	,941(**)	1	,161(**)	,004	-,103(**)	-,144(**)	,003	,242(**)
	Sig. (2-tailed)	,000	,000	,000	,822	,000	,000	,885	,000
TP_Contrib	Pearson Correlation	,179(**)	,161(**)	1	,057(**)	,209(**)	,268(**)	,004	-,030
	Sig. (2-tailed)	,000	,000	,000	,003	,000	,000	,856	,129
QT_DEP_IR	Pearson Correlation	-,006	,004	,057(**)	1	,031	,033	,007	,006
	Sig. (2-tailed)	,772	,822	,003	,000	,112	,091	,735	,742
QT_SM_RMI	Pearson Correlation	-,162(**)	-,103(**)	,209(**)	,031	1	,709(**)	,050(*)	-,237(**)
	Sig. (2-tailed)	,000	,000	,000	,112	,000	,000	,011	,000
VL_MR_ATU	Pearson Correlation	-,034	-,144(**)	,268(**)	,033	,709(**)	1	,102(**)	-,072(**)
	Sig. (2-tailed)	,081	,000	,000	,091	,000	,000	,000	,000
QT_REVISAO_RMI	Pearson Correlation	,000	,003	,004	,007	,050(*)	,102(**)	1	,197(**)
	Sig. (2-tailed)	,986	,885	,856	,735	,011	,000	,000	,000
Meses_Receb	Pearson Correlation	,194(**)	,242(**)	-,030	,006	-,237(**)	-,072(**)	,197(**)	1
	Sig. (2-tailed)	,000	,000	,129	,742	,000	,000	,000	,000

Após a identificação da alta correlação entre as características ID_SEG_2008 e ID_SEG_DIB, que representam respectivamente, a idade do segurado no ano de 2008 (ID_SEG_2008) que corresponde à data do estudo (janeiro/2008) e a idade do segurado no momento do início do benefício (ID_SEG_DIB), foi dirigida aos especialistas do INSS e DATAPREV a correlação já descrita e estes definiram que a característica ID_SEG_2008 deveria ser retirada do estudo, por tratar-se de informação cronológica e portanto derivada de outras informações cronológicas e demográficas dos segurados, sendo possível a sua recuperação em momentos

futuros, caso necessário. O quadro 6 apresenta a retirada da variável ID_SEG_2008 do modelo e o resultado de um novo teste de correlação, executado no software estatístico.

Quadro 6 - Variáveis métricas após análise de correlação (multicolinearidade)

		ID_SEG_DIB	TP_Contrib	QT_DEP_IR	QT_SM_RMI	VL_MR_ATU	QT_REVISAO_RMI	Meses_Receb
ID_SEG_DIB	Pearson Correlation	1	,179(**)	-0,006	-,162(**)	-0,034	0	0,194(**)
	Sig. (2-tailed)		0	0,772	0	0,081	0,986	0
TP_Contrib	Pearson Correlation	,179(**)	1	,057(**)	,209(**)	,268(**)	0,004	-0,03
	Sig. (2-tailed)	0		0,003	0	0	0,856	0,129
QT_DEP_IR	Pearson Correlation	-0,006	,057(**)	1	0,031	0,033	0,007	0,006
	Sig. (2-tailed)	0,772	0,003		0,112	0,091	0,735	0,742
QT_SM_RMI	Pearson Correlation	-,162(**)	,209(**)	0,031	1	,709(**)	,050(*)	-,237(**)
	Sig. (2-tailed)	0	0	0,112		0	0,011	0
VL_MR_ATU	Pearson Correlation	-0,034	,268(**)	0,033	,709(**)	1	0,102(**)	-,072(**)
	Sig. (2-tailed)	0,081	0	0,091	0		0	0
QT_REVISAO_RMI	Pearson Correlation	0	0,004	0,007	,050(*)	,102(**)	1	,197(**)
	Sig. (2-tailed)	0,986	0,856	0,735	0,011	0		0
Meses_Receb	Pearson Correlation	0,194(**)	-0,03	0,006	-,237(**)	-,072(**)	0,197(**)	1
	Sig. (2-tailed)	0	0,129	0,742	0	0	0	

A técnica de regressão logística é afetada por categorias de alta dispersão de seus elementos contendo pequenos quantitativos de elementos, influenciando no modelo e gerando grande quantidade de variáveis *dummies*, reduzindo a eficiência do modelo e no poder de predição do mesmo (HAIR et al., 2005; CORRAR, 2007).

Tomou-se a decisão de promover agregações nas categorias de natureza não-métrica, utilizou-se o critério de agregação para categorias que não ultrapassaram 100 ocorrências, na amostra de 2.622 benefícios, sendo portanto agrupadas em categoria definida neste estudo como “OUTROS” e não há qualquer

associação às categorias utilizadas no âmbito do INSS e DATAPREV em relação aos seus significados e descrições, segundo o APÊNDICE C.

A tabela 19 apresenta o resultado do teste Kolmogorov-Smirnov e Shapiro-Wilk, utilizado para verificar a distribuição das variáveis explicativas métricas, uma vez que para um processo de análise multivariada, neste caso a *logit*, a distribuição normal das variáveis é uma suposição imprescindível e deve ser verificada através de um teste não-paramétrico. A suposição de H_0 (a distribuição é normal) e H_1 (a distribuição não é normal) foi verificada através do teste Kolmogorov-Smirnov com verificação de *Lilliefors* (teste mais preciso e robusto), onde variáveis com significância menor que 5% ($sig < 0,05$), possuem distribuição normal, sendo descartada H_1 e aceita H_0 (CORRAR, 2007, p. 45).

Tabela 19 - Teste de normalidade Kolmogorov-Smirnov

	Kolmogorov-Smirnov(a)			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
ID_SEG_DIB	0,036	2613	0,000	0,992	2613	0,000
QT_DEP_IR	0,517	2613	0,000	0,029	2613	0,000
QT_SM_RMI	0,194	2613	0,000	0,765	2613	0,000
VL_MR_ATU	0,136	2613	0,000	0,848	2613	0,000
QT_REVISAO_RMI	0,530	2613	0,000	0,268	2613	0,000
Anos_Recebidos	0,214	2613	0,000	0,490	2613	0,000

a Lilliefors Significance Correction

Antes da geração do modelo logit através do *software* estatístico, foi executado o processo de separação da amostra em dois conjuntos com respectivamente 60% e 40% (*no software data, select cases, Random sample os cases, sample, 60%*) das ocorrências, este processo é identificado por Hosmer e Lemeshow (1989) como *amostra de treinamento* (60%) e *amostra de teste* (40%), enquanto que para Hair et al. (2005) estas amostras são denominadas como *Amostra de Análise* (60%) e *Amostra de Validação* (40%), para Corrar et al. (2007) as amostras são tratadas conforme a denominação de Hosmer e Lemeshow (1989).

Na execução do processo de regressão logística (*logit*) foram consideradas todas as variáveis métricas (exceto ID_SEG_2008, retirada em função de altos níveis de correlação) e todas as variáveis categóricas identificadas pelos especialistas após o processamento dos 127 questionários respondidos e classificados como corretos a partir do julgamento dos respondentes (juizes) (COOPER et SCHINDLER, 2003).

As variáveis categóricas utilizadas no modelo foram as resultantes de um agrupamento que considerou a categoria “OUTROS” como concentradora de todas as demais categorias que contivessem menos de 100 observações na amostra de 2.622 benefícios por incapacidade constituída para o estudo. O demonstrativo das agregações categóricas encontra-se no APÊNDICE D.

Na tabela 20, consta o sumário das informações processadas na geração do modelo, onde foram considerados 1588 observações na amostra de treinamento, sobre as quais foi gerado o modelo e 1030 observações relativas à amostra de teste, sobre a qual o modelo foi validado.

Tabela 20 - Sumário do processamento do modelo

Unweighted Cases(a)		N	Percent
Amostra de Treinamento	Incluído na análise	1.588	60,60%
	Missing Cases	4	0,20%
	Total	1.592	60,70%
Amostra de Testes		1.030	39,30%
Total		2.622	100,00%

a If weight is in effect, see classification table for the total number of cases.

A classificação da variável dicotômica correspondente à situação dos benefícios, foi analisada como 0 (zero) para os benefícios ativos e 1 (um) para os benefícios considerados como irregulares pelos Órgãos de controle e cessados por irregularidade, a partir da determinação destes órgãos.

A classificação inicial, sem a utilização das variáveis independentes (explicativas) consta na tabela 21, referindo-se somente aos valores do intercepto.

Na amostra de treinamento o modelo consegue classificar corretamente os 819 benefícios na situação de cessados corretamente, porém classifica incorretamente os 769 benefícios na situação de ativos como cessados, atingindo um percentual de classificação de 51,57% de correção. Em relação à amostra de testes, todos os benefícios foram classificados como ativos, embora 492 destes tenham sido cessados pelos Órgãos de controle, correspondendo ao percentual de classificação correto de 48%.

Tabela 21 - Tabela de classificação somente com a constante

			Amostra de Treinamento			Amostra de Testes		
			STATUS		Percentual Correto	STATUS		Percentual Correto
			ATIVO	CESSADO		ATIVO	CESSADO	
Step 0	STATUS	ATIVO	0	769	0	0	533	0
		CESSADO	0	819	100	0	492	100
	Percentual Final				51,57			48,00

As variáveis consideradas no modelo somente com o valor da constante (intercepto) são apresentadas na tabela 22.

Tabela 22 - Variáveis na equação com a constante

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	0,063	0,050	1,574	1	0,210	1,065

Através da execução do teste através de stepwise com o método *Likelihood value*, onde as variáveis independentes serão incluídas gradualmente e retiradas em função do nível de influência na probabilidade (*likelihood value*) sobre a relação entre os valores examinados e os previstos através da estatística qui-quadrado e da verossimilhança, considerando os níveis definidos para aceitação dos valores (5%) e de retirada dos mesmos (10%), considerando uma distribuição normal com 2 desvios-padrão (-1,96 a +1,96 sobre a curva normal).

Segundo Corrar et al. (2007, p. 294), o *likelihood value* é uma das principais medidas de avaliação do modelo *logit*, permitindo estimar a probabilidade de um evento ocorrer, através da estimativa de ocorrência do evento a partir de uma distribuição qui-quadrado, onde o evento ocorre quando o valor é 1, sendo neste caso o valor de $-2LL$ igual a zero. Também para Hair et al. (2005) e Pereira e Ness-Junior (2008), o *likelihood value* é uma medida que não possui significado intrínseco, sendo utilizada a diferença entre o valor do *likelihood value* do modelo somente com a constante e o do modelo final contendo as variáveis preditoras, esta diferença, quanto maior, atribuirá um maior potencial de previsão ao modelo, baseado nos coeficientes e nas variáveis independentes (VI) que integrarem o modelo final.

Tabela 23 - Teste de avaliação sequencial dos coeficientes

		Chi-square	df	Sig.
Step 1	Step	792,552	3	,000
	Block	792,552	3	,000
	Model	792,552	3	,000
Step 2	Step	279,165	4	,000
	Block	1071,716	7	,000
	Model	1071,716	7	,000
Step 3	Step	490,518	1	,000
	Block	1562,234	8	,000
	Model	1562,234	8	,000
Step 4	Step	102,567	4	,000
	Block	1664,802	12	,000
	Model	1664,802	12	,000
Step 5	Step	39,285	3	,000
	Block	1704,086	15	,000
	Model	1704,086	15	,000
Step 6	Step	26,209	1	,000
	Block	1730,295	16	,000
	Model	1730,295	16	,000
Step 7	Step	26,355	4	,000
	Block	1756,650	20	,000
	Model	1756,650	20	,000
Step 8	Step	13,116	1	,000
	Block	1769,766	21	,000
	Model	1769,766	21	,000
Step 9	Step	70,220	1	,000
	Block	1839,986	22	,000
	Model	1839,986	22	,000
Step 10	Step	9,381	1	,002
	Block	1849,367	23	,000
	Model	1849,367	23	,000

O teste dos coeficientes em 10 etapas via *stepwise*, chegou à estimativa constante na tabela 23, onde destaca-se a diferença entre o valor do qui-quadrado (chi-square) do passo 1 igual a 792,552 e o valor do qui-quadrado no passo 10, igual a 1849,367, com significância menor que 0,05, além da evolução do qui-quadrado relativo aos vários passos (*steps*), variando de 792,552 (passo 1) até 9,391 (passo 10). Segundo Corrar et al. (2007, p. 295), o principal objetivo do

likelihood value é verificar se a regressão é estatisticamente significativa e facilitar a comparação entre modelos alternativos.

No tabela 24 encontram-se demonstrados os coeficientes pseudo- r^2 de Cox e Snell (0,688) e o pseudo- r^2 Nagelkerke (0,918). O nível de significância do modelo está retratado pelos coeficientes pseudo- r^2 tanto o de Cox e Snell que possui o seu melhor valor próximo a 1 (um), mesmo sem poder atingi-lo. Já para o coeficiente pseudo- r^2 de Nagelkerke, pode atingir o valor máximo igual a 1 (um) (HAIR et al, 2005;CORRAR et al., 2007).

Considerando o níveis atingidos, o modelo possui alto poder preditivo em relação à amostra utilizada, uma vez que o r^2 de Cox e Snell consegue explicar 68,8% do log da razão de chance, enquanto o r^2 Nagelkerke consegue explicar 91,8% do log da razão de chance, ambos após a inclusão das variáveis explicativas independentes no modelo, através do processo *stepwise*, considerando como referência o modelo contendo apenas a constante (CORRAR et al., 2007, p. 308).

Ainda segundo Corrar et al. (2007), a tabela 23 apresenta considerável redução no coeficiente pseudo $-r^2$ loglikelihood (-2LL), mais uma característica que conduz à melhoria do ajuste do modelo após a inclusão das variáveis explicativas.

Tabela 24 - Resumo do modelo

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1407,309(a)	,393	,524
2	1128,145(b)	,491	,655
3	637,626(b)	,626	,835
4	535,059(b)	,649	,866
5	495,774(b)	,658	,878
6	469,565(b)	,664	,885
7	443,211(b)	,669	,893
8	430,095(b)	,672	,896
9	359,875(b)	,686	,915
10	350,494(b)	,688	,918

De acordo com Corrar et al. (2007), a tabela 25, fornecida pelo software estatístico, retrata o teste de Hosmer e Lemeshow, que consiste na separação da amostra de treinamento em 10 sub-amostras e comparação destas sub-amostras em relação ao poder preditivo do modelo, avaliando a sua correção em função da estatística qui-quadrado (valores existentes e valores preditos) nas situações de *ativos* e *cessados* dos benefícios da amostra de treinamento, apresenta alta significância (0,000 em relação a 0,05) e redução significativa do valor da estatística chi-square (qui-quadrado), quando o modelo foi gerado somente com a constante (sem as variáveis independentes).

Ainda segundo o autor, o testes Hosmer e Lemeshow consiste na comparação entre as freqüências preditas e as observadas, considerando o nível de significância (0,000), com 8 graus de liberdade, pode-se aceitar a hipótese de que não há diferenças significativas entre as classificações preditas e observadas, aceitando-se que o modelo representa a realidade de forma satisfatória.

Tabela 25 - Resultado do teste de Hosmer e Lemeshow

Step	Chi-square	df	Sig.
1	0,000	2	1,000
2	49,482	7	0,000
3	322,214	8	0,000
4	127,310	8	0,000
5	266,390	8	0,000
6	109,411	8	0,000
7	89,896	8	0,000
8	34,997	8	0,000
9	58,855	8	0,000
10	69,110	8	0,000

A tabela 26, apresenta o resultado do teste de Hosmer e Lemeshow sobre a amostra de treinamento, constituída por 60% da amostra de 2.622 benefícios por incapacidade, com 8 graus de liberdade (*degrees-of-freedom*).

Tabela 26 – Tabela de contingência para o teste Hosmer e Lemeshow

		STATUS = ATIVO		STATUS = CESSADO		Total
		Observed	Expected	Observed	Expected	
Step 10	1	158	158,976923	1	0,023077048	159
	2	157	158,8096831	2	0,190316864	159
	3	157	157,6908544	2	1,309145579	159
	4	151	152,7576755	8	6,242324514	159
	5	132	118,6172961	27	40,38270392	159
	6	14	21,45031603	145	137,549684	159
	7	0	0,656239805	159	158,3437602	159
	8	0	0,038886921	159	158,9611131	159
	9	0	0,002114475	158	157,9978855	158
	10	0	1,08577E-05	158	157,9999891	158

A tabela 27, apresenta o resultado da classificação dos casos constantes na amostra de treinamento (60%) e na amostra de teste (40%) do modelo.

Tabela 27 - Classificação final do modelo

			Casos selecionados (Treinamento)			Casos não-selecionados (Teste)		
			STATUS		Percentual Correto	STATUS		Percentual Correto
			ATIVO	CESSADO		ATIVO	CESSADO	
Passo 10	STATUS	ATIVO	747	22	97,14	524	9	98,31
		CESSADO	34	785	95,85	25	467	94,92
		Percentual geral			96,47			96,68

A partir dos resultados dos testes de Hosmer e Lemeshow, buscando os níveis de ajustamento na amostra de teste e das relações entre observações constantes na amostra e os valores esperados a partir do modelo gerado, foi feita a comparação entre a aplicação do modelo gerado na amostra de treinamento sobre a amostra de teste, correspondente à 40% da amostra de 2.622 benefícios por incapacidade que compõe o modelo.

Em relação ao observado no tabela 27, o modelo gerado é capaz de classificar os benefícios da amostra de validação com 96,47% de precisão em relação à sua situação (ativo ou cessado), representada pela variável dicotômica STATUS, enquanto na amostra de teste o nível de precisão obtido foi de 96,68%,

segundo Hair et al. (2005), o nível de predição do modelo deverá ser o obtido na classificação da amostra de teste, definidas segundo Hair et al. (2005) como amostra de análise (60%) e amostra de validação (40%).

Na amostra de análise (60% do total) o erro do tipo I (benefícios cessados que foram classificados como ativos) foi de 4,15% (34 casos) e o erro do tipo II (benefícios ativos que foram classificados como cessados) foi de 2,86% (22 casos).

Na amostra de validação (40% do total) o erro do tipo I foi de 5,08% (25 casos) e o erro do tipo II foi de 1,69% (9 casos).

Diante dos custos de classificação descritos acima, observa-se que o modelo não penaliza os segurados que possuem benefícios ativos, já que o percentual de classificações de ativos como cessados (erros tipo II) é proporcionalmente muito inferior ao erro tipo I, tanto na amostra de análise quanto na amostra de validação.

O modelo final apresenta 4 variáveis métricas e 7 categóricas, segundo a tabela 27. Em sua fase inicial o estudo possuía 20 variáveis independentes, o método *stepwise (likelihood ratio)* utilizado na geração do modelo através do software estatístico, desconsiderou 9 (nove) variáveis a partir das significâncias definidas para entrada (5%) e para eliminação das variáveis (10%).

Segundo Corrar et al. (2007, p. 286) a equação que identifica a probabilidade associada à ocorrência de determinado fenômeno, neste caso o *score* que foi atribuído ao benefício por incapacidade, é definido como:

$$P(\text{evento}) = \frac{1}{1 + e^{-(b_0 + b_1x_1 + \dots + b_nx_n)}}$$

Onde: P(evento) é a probabilidade de ocorrência do evento

e - logaritmo natural

b_0 a b_n – Coeficientes obtidos na geração do modelo logit

x_1 a x_n – Variáveis explicativas obtidas após definição do modelo logit

As variáveis selecionadas na tabela 28 foram identificadas através do processo *stepwise* do software estatístico e nos passos onde cada variável categórica foi incluída no modelo, estas possuíam significância menor que 5% (sig < 0,05) e mais de um grau de liberdade, em pelo menos uma de suas categorias. As variáveis métricas, possuem a significância de 5% (sig < 0,05) e um grau de liberdade.

Segundo Sincich (1996), no modelo de regressão logística a influência das variáveis independentes pode ser obtida através do antilog do coeficiente β_i (e^{β_i}) de cada variável independente. O valor de e^{β_i} representa a mudança nas chances de ser observado o evento previsto pelo modelo ($Y = 1$ (benefício incorreto)), quando a variável métrica independente, associada a este coeficiente é elevada em uma unidade, mantendo-se as outras variáveis constantes.

Ainda segundo Sincich (1996), a expressão $e^{\beta_i} - 1$ indica um percentual nas chances do evento alvo do modelo ser observado, alegando que um percentual negativo demonstra o quanto as chances do evento a ser previsto diminuem com o aumento de uma unidade na variável métrica independente associada a este coeficiente, mantendo-se as outras variáveis independentes constantes.

A partir da tabela 28, observa-se que a variável VL_MR_ATU (Valor recebido pelo segurado) a cada unidade acrescentada, reduz em 0,4% a chance de um benefício tornar-se cessado. A variável QT_REVISAO_RMI (Quantidade de revisões solicitadas sobre quantidade de salários-mínimos no início do benefício) a cada acréscimo unitário, reduz em 76,15% as chances de um benefício assumir o status de cessado. A variável Anos_Recebidos (Tempo de recebimento do benefício em anos) reduz em 61,20% as chances de um benefício assumir o status de cessado, a cada ano de recebimento do benefício.

Ainda segundo a tabela 28, pode-se afirmar que os benefícios relativos à espécie de AUXILIO DOENÇA PREVIDENCIÁRIO (CS_ESPECIE = 31) possuem o probabilidade de tornarem-se incorretos (cessados), os segurados filiados à Previdência Social através da categoria DESEMPREGADO, também apresentam a maior probabilidade, assim como as atividades relativas ao setor INDUSTRIA e os segurados que formalizaram o último vínculo empregatício na Região Nordeste. Também possuem as maiores chances de tornar-se irregulares (cessados) os benefícios concedidos através de CONCESSÃO DESCENTRALIZADA ou da CONCESSÃO PRISMA/SUB.

Tabela 28 - Variáveis independentes incorporadas ao modelo

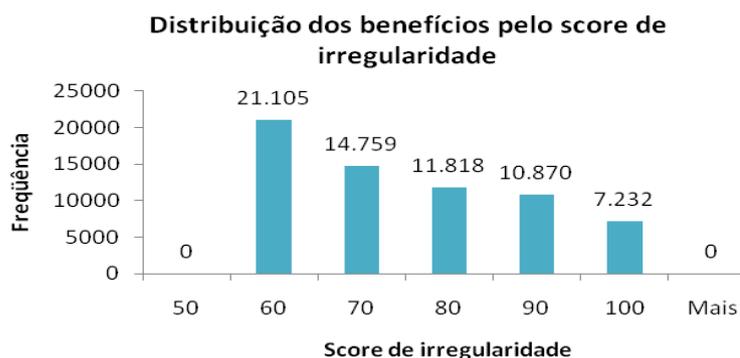
		B	S.E.	Wald	df	Sig.	Exp(B)	95,0% C.I. for EXP(B)	
Variável explicativa	Significado da variável							Lower	Upper
CS_ESPECIE				39,67119	3	1,25E-08			
CS_ESPECIE(1)	AUXILIO DOENCA PREVIDENCIARIO	1,12253	1,3579	0,683339	1	0,4084394	3,07263	0,21460	43,99396
CS_ESPECIE(2)	APOSENTADORIA INVALIDEZ PREVIDENCIARIA	-1,2853	1,343	0,915997	1	0,3385284	0,27656	0,01989	3,84527
CS_ESPECIE(3)	AUXILIO DOENCA POR ACIDENTE DO TRABALHO	0,15368	1,6481	0,008696	1	0,925704	1,16612	0,04612	29,48371
CS_FORMA_FILIACAO_GRUPO				42,83119	4	1,12E-08			
CS_FORMA_FILIACAO_GRUPO(1)	DESEMPREGADO	1,97033	0,6483	9,238021	1	0,0023704	7,17303	2,01327	25,55653
CS_FORMA_FILIACAO_GRUPO(2)	EMPREGADO	-0,134	0,6746	0,039471	1	0,8425182	0,87457	0,23311	3,28114
CS_FORMA_FILIACAO_GRUPO(3)	SEGURADO ESPECIAL	0,73353	1,449	0,256273	1	0,6126924	2,08242	0,12167	35,64141
CS_FORMA_FILIACAO_GRUPO(4)	AUTONOMO	0,00207	0,6725	9,43E-06	1	0,9975494	1,00207	0,26819	3,74409
QT_SM_RMI		0,69685	0,0811	73,86995	1	8,34E-18	2,00742	1,71247	2,35316
VL_MR_ATU		-0,0035	0,0005	59,27295	1	1,37E-14	0,99648	0,99558	0,99737
CS_RAMO_ATIVIDADE_GRUPO				18,63562	3	0,0003252			
CS_RAMO_ATIVIDADE_GRUPO(1)	COMERCARIO	-1,2341	0,8603	2,05786	1	0,1514227	0,29111	0,05393	1,57147
CS_RAMO_ATIVIDADE_GRUPO(2)	INDUSTRIARIO	1,43037	0,9808	2,126811	1	0,1447416	4,18024	0,61142	28,58027
CS_RAMO_ATIVIDADE_GRUPO(3)	RURAL	-2,4336	1,42	2,9373	1	0,0865553	0,08772	0,00543	1,41831
QT_REVISAO_RMI		-1,4335	0,5296	7,326824	1	0,0067933	0,23848	0,08446	0,67333
UnidadeTerritorial				16,15	4	0,0028243			
UnidadeTerritorial(1)	REGIÃO NORTE	0,66605	0,9927	0,450199	1	0,5022407	1,94652	0,27817	13,62120
UnidadeTerritorial(2)	REGIÃO NORDESTE	1,28802	0,8309	2,403188	1	0,1210883	3,62561	0,71146	18,47625
UnidadeTerritorial(3)	REGIÃO SUDESTE	0,32232	0,7942	0,164699	1	0,684867	1,38032	0,29104	6,54656
UnidadeTerritorial(4)	REGIÃO SUL	-1,0015	0,9	1,238307	1	0,2657975	0,36732	0,06294	2,14356
CS_SISTEMA_CONCES_GRUPO				61,73715	4	1,25E-12			
CS_SISTEMA_CONCES_GRUPO(1)	CONCESSAO PRISMA/SUB	4,00699	1,1938	11,26673	1	0,0007891	54,9813	5,29759	570,62535
CS_SISTEMA_CONCES_GRUPO(2)	CONCESSAO ON-LINE	-1,8257	1,005	3,299955	1	0,0692818	0,1611	0,02247	1,15501
CS_SISTEMA_CONCES_GRUPO(3)	CONCESSAO DESCENTRALIZADA	22,1688	2489,9	7,93E-05	1	0,9928962	4,20E+09	0,00000	
CS_SISTEMA_CONCES_GRUPO(4)	CONCESSAO ON-LINE, SEM INDICE	-1,1641	1,0724	1,17834	1	0,2776943	0,31221	0,03816	2,55421
CS_BATIMENTO_CNIS_GRUPO(1)	NAO HOUVE UTILIZACAO DE DADOS DO CNIS	1,4744	0,3767	15,31766	1	9,09E-05	4,36841	2,08765	9,14088
Anos_Recebidos		-0,9468	0,0828	130,8768	1	2,63E-30	0,38798	0,32989	0,45631
Constant		2,9777	2,0283	2,15528	1	0,1420805	19,6427		

4.2 Aplicação do modelo gerado

Após a aplicação do modelo gerado no universo dos benefícios por incapacidade na situação de ativo em dezembro de 2007 (1.399.296 de recebedores de benefícios por incapacidade), através de uma ferramenta de tecnologia da informação construída na linguagem Microsoft Visual Basic versão 6.0 com service Pack 6.0 para sistema operacional Windows. Foram identificados 65.784 benefícios que encontravam-se como ativos e que segundo o modelo, deveriam ser analisados detalhadamente, devido ao grau de similaridade que apresentaram em relação aos benefícios cessados por irregularidade, em decorrência das características avaliadas no modelo.

Conforme observado na figura 5, há 7.232 benefícios com probabilidade (*score*) entre 90% e 100% de similaridade com um benefício por incapacidade cessado.

Figura 5 - Histograma dos benefícios com score entre 50% e 100%

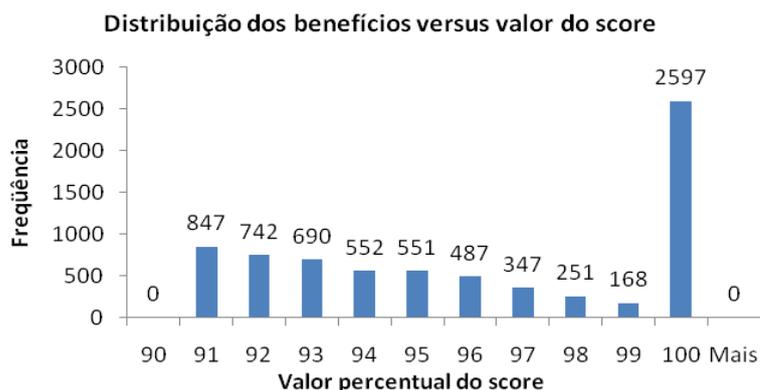


A figura 5, apresenta a análise das variáveis independentes identificadas no modelo de Regressão Logística relativa aos 65.784 benefícios selecionados com percentual maior que 50%, após aplicação do modelo gerado sobre o cadastro de benefícios. Foi identificada a situação em que os benefícios são apontados como

cessados, segundo o modelo que utilizou os benefícios cessados pelos Órgãos de Controle e uma amostra probabilística sistemática dos benefícios por incapacidade, embora estes benefícios sejam considerados em situação regular (*ativos*) a partir do cadastro de benefícios mantido pela DATAPREV e sob gestão do Ministério da Previdência e Assistência Social (MPAS).

Sobre o modelo *logit* resultante, são apresentadas as distribuições das variáveis independentes (*Vis*) que exercem influência no modelo e portanto foram selecionados pela ferramenta estatística. A figura 6 apresenta a distribuição dos benefícios com *score*, resultante da aplicação do modelo, entre 90% e 100%.

Figura 6 - Distribuição do benefícios na faixa de risco entre 90% e 100%



A seguir são apresentados os valores gerados para pagamento dos benefícios identificados pelo modelo com *score* superior a 50%. Os valores foram calculados a partir da solução de tecnologia, desenvolvida para esta finalidade. Foi considerada a quantidade de salários-mínimos recebida pelo segurado em cada competência, convertida para o valor do salário-mínimo em Reais (R\$) nas respectivas competências. A soma dos valores correspondentes ao recebimento de cada benefício, valores gerados pela Previdência Social para pagamento, é apresentada através do agrupamento das variáveis explicativas identificadas pelo modelo de regressão logística gerado.

São apresentados os valores para os benefícios que possuem scores entre 90% e 100%, as faixas de 50% a 100% constam no APÊNDICE E e os valores relativos às faixas de scores entre 99% e 100% constam no APÊNDICE F.

Tabela 29 - Distribuição dos Benefícios por Espécie – score entre 50% e 100%

Especie	Descrição	Quantidade	Valor gerado (R\$)	% do Valor gerado
31	AUXILIO DOENCA PREVIDENCIARIO	53.070	958.028.602,87	71,75%
32	APOSENTADORIA INVALIDEZ PREVIDENCIARIA	10.726	299.110.452,61	22,40%
91	AUXILIO DOENCA POR ACIDENTE DO TRABALHO	804	25.031.575,66	1,87%
92	APOSENT. INVALIDEZ ACIDENTE TRABALHO	1.184	52.982.047,44	3,97%
Total		65.784	1.335.152.678,58	

Tabela 30 -Distribuição dos Benefícios por Espécie - score entre 90% e 100%

Especie	Descrição	Quantidade	Valor gerado (R\$)	% do Valor gerado
31	AUXILIO DOENCA PREVIDENCIARIO	5.138	95.208.874,67	54,65%
32	APOSENTADORIA INVALIDEZ PREVIDENCIARIA	1.601	50.923.955,10	29,23%
91	AUXILIO DOENCA POR ACIDENTE DO TRABALHO	315	16.787.857,37	9,64%
92	APOSENT. INVALIDEZ ACIDENTE TRABALHO	178	11.305.224,10	6,49%
Total		7.232	174.225.911,24	

Tabela 31 - Distribuição dos Benefícios por Espécie score entre 99% e 100%

Espécie	Descrição	Quantidade	Valor gerado (R\$)	% do Valor gerado
31	AUXILIO DOENCA PREVIDENCIARIO	1.015	45.188.806,61	41,09%
32	APOSENTADORIA INVALIDEZ PREVIDENCIARIA	1.160	38.398.010,67	34,91%
91	AUXILIO DOENCA POR ACIDENTE DO TRABALHO	270	16.132.925,06	14,67%
92	APOSENT. INVALIDEZ ACIDENTE TRABALHO	152	10.267.195,93	9,33%
Total		2.597	109.986.938,27	

Tabela 32 - Distribuição por Sistema de Concessão – score entre 90% e 100%

Sistema de concessão	Descrição	Quantidade	Valor gerado (R\$)	% do Valor gerado
1	CONCESSAO FORMULARIO CCE	2	83.344,63	0,05%
4	CONCESSAO PRISMA/SUB	551	21.263.017,58	12,20%
5	CONCESSAO ON-LINE	2.931	27.805.446,70	15,96%
7	CONCESSAO DESCENTRALIZADA	2.519	108.559.312,34	62,31%
8	CONCESSAO ON-LINE, SEM INDICE	1.198	16.102.118,49	9,24%
11	CONCESSAO PELA INTERNET	31	412.671,50	0,24%
Total		7.232	174.225.911,24	

A figura 7 apresenta a distribuição dos benefícios selecionados pelo modelo a partir da quantidade de anos de recebimento do benefício (Anos_Recebidos) e na tabela 32, são apresentados os percentuais relativos à distribuição dos benefícios.

Figura 7 - Quantidade de benefícios por ano de recebimento

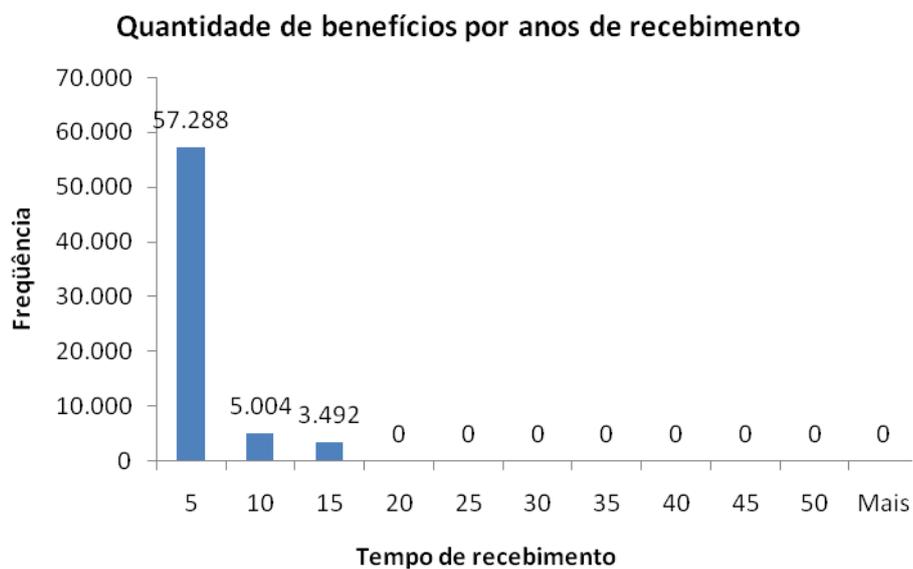


Tabela 33 - Distribuição dos benefícios por tempo de recebimento

Tempo de recebimento (em anos)	Quantidade de benefícios	Percentual
5	57.288	87,09%
10	5.004	7,61%
15	3.492	5,31%
20	0	0,00%
25	0	0,00%
30	0	0,00%
35	0	0,00%
40	0	0,00%
45	0	0,00%
50	0	0,00%
Total	65.784	

Tabela 34 - Benefícios por forma de filiação - score entre 90% e 100%

Forma de Filiação	Descrição	Quantidade	Valor gerado (R\$)	% do Valor gerado
0	DESEMPREGADO	4.738	80.243.523,81	46,06%
1	EMPREGADO	1.472	71.173.138,21	40,85%
2	TRABALHADOR AVULSO	31	331.378,40	0,19%
3	EMPRESARIO	18	805.215,35	0,46%
4	DOMESTICO	70	1.666.930,40	0,96%
5	FACULTATIVO	47	1.149.049,46	0,66%
6	EQUIPARADO A AUTONOMO	6	238.136,63	0,14%
7	SEGURADO ESPECIAL	487	10.718.639,08	6,15%
8	AUTONOMO	358	7.853.412,90	4,51%
9	OPTANTE PELA LEI 6.184/74	5	46.487,00	0,03%
Total		7.232	174.225.911,24	

Tabela 35 - Distribuição por Ramo de Atividade – score entre 90% e 100%

Ramo de Atividade	Descrição	Quantidade	Valor gerado (R\$)	% do Valor gerado
1	BANCARIO	37	4.263.673,59	2,45%
2	COMERCARIO	5.021	91.908.971,00	52,75%
3	TRANSPORTES E CARGA	110	4.832.042,47	2,77%
4	FERROVIARIO	3	266.860,92	0,15%
5	INDUSTRIARIO	1.139	48.806.908,86	28,01%
6	MARITIMO	1	94.605,44	0,05%
7	SERVIDOR PUBLICO	128	4.571.529,75	2,62%
8	RURAL	715	18.057.435,75	10,36%
9	IRRELEVANTE	78	1.423.883,46	0,82%
Total		7.232	174.225.911,24	

Tabela 36 - Distribuição por Acesso ao CNIS – score entre 90% e 100%

Tipo de Acesso	Descrição	Quantidade	Valor gerado (R\$)	% do Valor gerado
*	NAO HOUVE UTILIZACAO DE DADOS DO CNIS	12	795.595,09	0,46%
000	NAO HOUVE UTILIZACAO DE DADOS DO CNIS	7.001	170.319.369,90	97,76%
001	ALTERACAO NOS DADOS BASICOS	25	348.451,36	0,20%
010	ALTERACAO NOS VINCULOS	7	88.135,00	0,05%
011	ALTER DADOS BASICOS, ALTERACAO VINCULOS	25	382.204,84	0,22%
020	INCLUSAO NOS VINCULOS	10	168.933,02	0,10%
021	ALTER DADOS BASICOS, INCLUSAO VINCULOS	8	95.222,89	0,05%
030	INCLUSAO/ALTERACAO NOS VINCULOS	4	68.347,11	0,04%
031	ALTER DADOS BASICOS, INCL/ALTER VINCULOS	17	262.224,77	0,15%
100	ALTERACAO NAS REMUNERACOES	1	5.520,00	0,00%
101	ALTER DADOS BASICOS, ALTER REMUNERACOES	1	5.870,00	0,00%
120	INCLUSAO VINCULOS, ALTER REMUNERACOES	1	6.920,00	0,00%
201	ALTER DADOS BASICOS, INCL REMUNERACOES	2	25.507,37	0,01%
211	ALT DADOS BAS, ALT VINCULOS, INC REMUNER	4	46.690,10	0,03%
220	INCLUSAO VINCULOS, INCLUSAO REMUNERACOES	2	27.565,47	0,02%
221	ALT DADOS BAS, INCL VINCULOS, INCL REMUN	1	13.520,00	0,01%
500	HOUVE UTIL. DADOS DO CNIS, SEM ALTERACAO	111	1.565.834,32	0,90%
Total		7.232	174.225.911,24	

Tabela 37 – Distribuição por Região Geográfica – score entre 90% e 100%

Unidade Geográfica	Quantidade	Valor gerado (R\$)	% do Valor gerado
NÃO INFORMADA	1	32.123,37	0,02%
CENTRO-OESTE	156	3.478.445,49	2,00%
NORDESTE	2.382	42.125.464,77	24,18%
NORTE	375	6.526.157,40	3,75%
SUDESTE	3.756	101.701.320,59	58,37%
SUL	562	20.362.399,62	11,69%
Total	7.232	174.225.911,24	

A figura 8 apresenta a distribuição quantitativa dos benefícios identificados pelo modelo distribuídos pelo valor recebido na competência dez/2007, com scores entre 50% e 100% de similaridade com benefícios cessados, devido às ações dos órgãos de controle da Previdência Social.

Figura 8 - Distribuição dos benefícios pelo valor da Renda Mensal

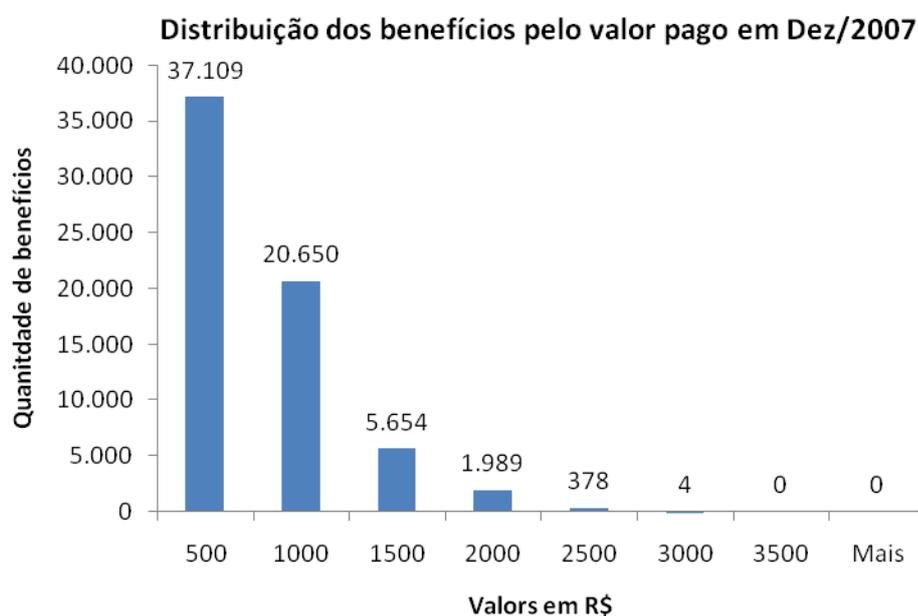


Tabela 38 - Distribuição dos segurados identificados por faixa de renda

Valor Recebido Mensalmente (em R\$)	Quantidade de benefícios	Percentual
500,00	37.109	56,41%
1.000,00	20.650	31,39%
1.500,00	5.654	8,59%
2.000,00	1.989	3,02%
2.500,00	378	0,57%
3.000,00	4	0,01%
3.500,00	0	0,00%
Mais	65.784	

A figura 9 apresenta a distribuição quantitativa dos benefícios identificados pelo modelo distribuídos a partir da quantidade de salários-mínimos recebida na competência dez/2007, com scores entre 50% e 100% de similaridade com benefícios cessados, devido às ações dos órgãos de controle da Previdência Social.

Figura 9 - Distribuição dos benefícios por Quantidade de salários-mínimos

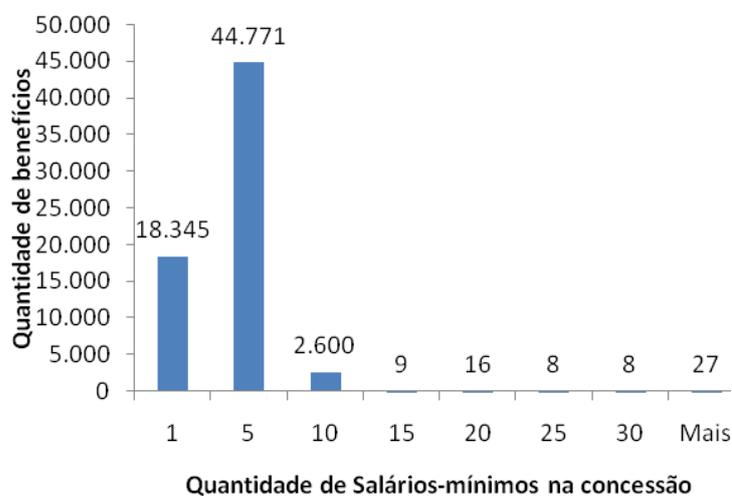


Tabela 39 - Distribuição dos segurados por valor de concessão

Quantidade de Salários-Mínimos	Quantidade de benefícios	Percentual
1	18.345	27,89%
5	44.771	68,06%
10	2.600	3,95%
15	9	0,01%
20	16	0,02%
25	8	0,01%
30	8	0,01%
Mais	27	0,04%
Total	65.784	

Tabela 40 - Distribuição pela Quantidade de revisões – score entre 90% e 100%

Quantidade de Revisões	Quantidade de benefícios	Valor gerado (R\$)	% do Valor gerado
0	6.924	160.213.733,57	91,96%
1	240	10.135.473,43	5,82%
2	47	2.639.701,89	1,52%
3	9	445.723,62	0,26%
4	7	492.896,83	0,28%
5	1	26.588,00	0,02%
6	2	127.214,54	0,07%
7	1	70.783,44	0,04%
8	1	73.795,92	0,04%
Total	7.232	174.225.911,24	

O quadro 7, apresenta as variáveis explicativas selecionadas pelo modelo e os passos (steps) nos quais foram incluídas através do processo stepwise, definido no software estatístico.

Quadro 7 - Variáveis consideradas no modelo final

A variável foi incorporada no passo 1: CS_ESPECIE.
A variável foi incorporada no passo 2: CS_SISTEMA_CONCES_GRUPO.
A variável foi incorporada no passo 3: Anos_Recebidos.
A variável foi incorporada no passo 4: CS_FORMA_FILIACAO_GRUPO.
A variável foi incorporada no passo 5: CS_RAMO_ATIVIDADE_GRUPO.
A variável foi incorporada no passo 6: CS_BATIMENTO_CNIS_GRUPO.
A variável foi incorporada no passo 7: UnidadeTerritorial.
A variável foi incorporada no passo 8: VL_MR_ATU.
A variável foi incorporada no passo 9: QT_SM_RMI.
A variável foi incorporada no passo 10: QT_REVISAO_RMI.

5 CONCLUSÃO

5.1 Análise dos resultados

A pesquisa sobre a conformidade dos benefícios por incapacidade baseou-se nas opiniões dos funcionários especialistas da DATAPREV e INSS, elegendo-se um conjunto de características que foi confirmado através de um *survey* de experiência, identificando-se 20 características que foram submetidas ao modelo de regressão logística tendo como referência um conjunto de 1.211 benefícios por incapacidade cessados por irregularidade, a partir de um conjunto de 14.286.545 de benefícios, segundo os órgãos de controle da Previdência Social. A amostra probabilística sistemática contendo 1.211 benefícios acidentários na situação de ativos foi obtida a partir do conjunto 2.373.704 de benefícios, sendo considerados 1.420.673 como integrantes do universo de benefícios por incapacidade ativos.

O modelo logístico resultante do estudo, considerou 10 variáveis explicativas independentes e conseguiu classificar com 96,68% de precisão a situação dos benefícios como ativos ou cessados (decorrentes de irregularidade), selecionando 65.784 benefícios com níveis de *score* entre 50% e 100%, que totalizam o valor de R\$ 1.335.152.678,58 disponibilizados para pagamento ao longo da vigência destes benefícios, a partir de um universo de 1.420.673 benefícios ativos na competência dezembro/2007, correspondendo o nível de acerto a 4,63% do universo na referida competência.

O nível de precisão na classificação dos quantitativos identificados e os valores monetários apresentados, são evidências de que há critérios comuns entre os benefícios cessados por irregularidade e os benefícios ativos, capazes de

propiciar a verificação de conformidade dos benefícios ativos a partir das características resultantes da análise dos órgãos de controle da Previdência Social e que este modelo multivariado, se utilizado nos processos de concessão e manutenção de benefícios por incapacidade, poderá evitar o pagamento indevido de 4,63% dos benefícios, além de auxiliar nas ações de auditoria contínua, possibilitando a análise de benefícios em momentos diferentes, a partir da análise por faixas de *score* atribuídas pelo modelo, podendo representar ações de recuperação de valores já pagos aos segurados ou evitar o pagamento indevido de valores financeiros da ordem de R\$ 1,335 bilhão.

O estudo apresentou o mapeamento de 10 características explicativas, a partir de um universo de 20 variáveis independentes, tendo sido identificadas 4 variáveis métricas e 6 taxonômicas, resultando em uma variável dependente (*score*) com valores entre 0,50 e 1,00 com referencial de corte (*cut-off*) de 0,50, classificando como benefícios incorretos àqueles que apresentaram *score* superior a 0,50.

Dentre as variáveis explicativas taxonômicas foram identificadas características que poderão ser incorporadas aos sistemas de tecnologia da informação da Previdência Social, contribuindo para a redução dos riscos de concessão ou pagamentos indevidos aos segurados na área de benefícios por incapacidade. Os maiores riscos identificados são relativos aos benefícios da categoria de Auxílio Doença Previdenciário (espécie 31), detentores de vínculo como desempregado, que desenvolvam atividades no setor industrial, que exerçam atividade na Região Nordeste, que tenha sido concedidos através dos sistemas de concessão descentralizada ou do sistema de concessão Prisma/SUB e cujas informações

demográficas ou contributivas não tenham sido confirmadas pelo Cadastro Nacional de Informações Sociais (CNIS) no momento da concessão do benefício.

Dentre as variáveis explicativas métricas identificadas pelo modelo, para cada alteração unitária ocorre uma redução ou aumento na chance de um benefício ter a sua situação alterada entre correto ou incorreto. Para a variável VL_MR_ATU (Valor recebido pelo segurado) a cada unidade acrescentada, ocorre a redução de 0,4% na chance de um benefício tornar-se cessado. A variável QT_REVISAO_RMI (Quantidade de revisões solicitadas sobre a quantidade de salários-mínimos no início do benefício), a cada acréscimo unitário, ocorre a redução de 76,15% das chances de um benefício assumir o status de cessado. Para a variável Anos_Recebidos (Tempo de recebimento do benefício em anos), ocorre a redução de 61,20% nas chances de um benefício por incapacidade assumir o status de cessado, a cada ano de recebimento do benefício.

As variáveis demográficas relativas aos segurados, não apresentaram nenhum poder preditivo no modelo resultante, sendo relevantes na geração do modelo apenas as características empregatícias e concessórias em poder da Previdência Social.

O modelo gerado permite que neste estudo seja desconsiderada a hipótese (H_0) de que não há critérios comuns nos benefícios por incapacidade cessados que possibilitem a verificação da conformidade legal dos benefícios por incapacidade concedidos e mantidos. O resultado do modelo assegura e confirma a aceitação da hipótese (H_1) que refere-se ao fato de que a análise das características dos benefícios cessados, possibilita a geração de Indicadores Antecedentes que serão aplicados na verificação de legitimidade dos benefícios mantidos pela Previdência Social.

O modelo gerado apresenta característica conservadora em suas previsões, uma vez que os erros de classificação mais significativos não penalizam os segurados, uma vez que foram identificados erros (tipo I) de 2,86% e 1,69% nos benefícios ativos que foram classificados como incorretos (cessados), enquanto foram identificados erros (tipo II) de 4,15% e 5,08% para os benefícios cessados (incorretos) que foram classificados como ativos, neste caso necessitando que outras ações sejam tomadas, uma vez que o nível de precisão do modelo não contempla estas situações.

O desenvolvimento de programas auxiliares foi de relevante importância na aplicação do modelo gerado e na estimativa dos valores recebidos pelos segurados, devido ao grande volume de observações (1,4 milhão), uma vez que os programas estatísticos utilizados não comportariam este volume de observações, inviabilizando o processo de verificação dos resultados.

Através da aplicação do modelo resultante do estudo, foi possível identificar que há 7.232 benefícios com scores entre 90% e 100%, resultando em R\$ 174.225.911,24, fato que evidencia a execução de avaliações detalhadas sobre estes benefícios, porém, também foram evidenciados 2.597 benefícios com scores entre 99% e 100%, resultando em R\$ 109.986.938,27 gerados para pagamento dos segurados, indicando que as ações de acompanhamento intensivo e avaliação detalhada deveriam ser aplicadas, prioritariamente, sobre este último conjunto de benefícios, em função dos valores envolvidos e do quantitativo evidenciado, seguindo o princípio da materialidade.

5.2 Considerações Finais

O modelo gerado mostrou-se adequado na mensuração dos impactos de sua aplicação sobre os benefícios das categorias acidentárias, mantidos e concedidos pela Previdência Social Brasileira, embasado em características comuns identificadas entre os benefícios cessados (encerrados por definição institucional) e os ativos (concedidos e pagos regularmente) no período analisado (jan/2007 a dez/2007). A identificação das características foi obtida através de um *survey de experiência* realizado com os técnicos da Previdência Social e DATAPREV, possibilitando a geração de indicadores antecedentes, embasados em Análises Projetivas, que poderão integrar os processos de validação e controle da Previdência Social, seja através de processos de negócios ou padrões de análise que serão desenvolvidos pelas equipes de Auditoria nos processos de verificação de conformidade, amparados pela geração e análise de *scores* de risco para cada benefício das categorias analisadas.

O estudo possibilitou o mapeamento das variáveis que influenciam na conformidade dos benefícios acidentários, com ênfase nas subcategorias de Auxílio por Acidente de Trabalho e Aposentadoria por Invalidez, possibilitando também, a geração de indicadores a partir das informações sob gestão da Previdência Social e DATAPREV, fundamentando as avaliações em profundidade dos processos de concessão e acompanhamento destas categorias de benefícios acidentários.

Através da geração do modelo de regressão logística, foi possível constatar a categorização dos benefícios acidentários em função do nível de conformidade com a legislação vigente e seus *scores* de classificação, resultantes da aplicação do modelo sobre cada benefício por incapacidade, tendo como base a análise das

características definidas pelos especialistas e dos resultados inferenciais obtidos pelo modelo *logit*.

Apoiado na Análise Projetiva, o modelo gerado confirma a possibilidade de realizar inferência sobre o comportamento futuro dos benefícios, apoiado em Indicadores Antecedentes, reafirmando a estabilidade da legislação sobre esta categoria de benefícios. A suposição de continuidade, relativa à presença das forças que moldaram o passado, são identificadas no presente e deverão influenciar o futuro (continuidade), garantindo estabilidade ao modelo e assegurando que as variáveis utilizadas na elaboração do modelo, bem como as informações oriundas dos especialistas, são fatores que embasam as projeções, considerando o conjunto das informações relativas ao período de junho/1994 a dez/2007 e gerenciadas pela Previdência Social Brasileira através da atuação da sua Empresa de Tecnologia, Informações e Comunicação (DATAPREV).

6 SUGESTÕES E RECOMENDAÇÕES

Este estudo utilizou informações relativas aos benefícios por incapacidade concedidos após julho/1994 devido à falta de completeza das informações para períodos anteriores, fato que poderia, em função da utilização da técnica de regressão logística, levar ao comprometimento da precisão do estudo devido à possível identificação de informações incompletas (*missing cases*) ou valores divergentes em relação ao universo analisado (*out liers*).

Como sugestão para novos estudos, fica a proposta de utilização da técnica de Redes Neurais Artificiais (RNA), técnica que não é influenciada pelas incorreções ou omissões das informações sob análise, podendo utilizar todo o acervo de benefícios sob gestão da Previdência Social na elaboração de modelos preditivos.

A ampliação do conjunto de características dos benefícios por incapacidade, poderá incorporar novos níveis de predição aos modelos, através da incorporação de outras características em suas análises, tais como, os laudos médicos decorrentes dos processos de concessão dos benefícios, além dos vínculos empregatícios e das alterações de exercício profissional ao longo da vida laborativa dos segurados.

7 BIBLIOGRAFIA

- AAKER, D.A et al. **Pesquisa de Marketing**. São Paulo:Atlas,2001.
- ABBAD, Gardênia; CARVALHO, Renata Silveira; ZERBINI, Thais. Evasão em Curso Via Internet: Explorando Variáveis Explicativas. **RAE – Eletrônica**, Rio de Janeiro, v. 5, n. 2, art. 17, jul/dez. 2006. Disponível em: WWW.rae.com.br/eletronica. Acessado em 15/04/2008.
- ALLEN, R.G.D. **The Assumptions of Linear Regression**. *Economica*, vol.6, n.22, p.191-201, 1939.
- ALMEIDA, F.C.; SIQUEIRA, J.O. Comparação entre Regressão Logística e Redes Neurais na previsão de falência de bancos Brasileiros. Terceiro Congresso Brasileiro de Redes Neurais. V. 4. Florianópolis, 1997. Disponível em: <http://www.fia.com.br/proinfo/artigos/Florianopolis.PDF> acessado em 11/07/2008.
- ALTMAN, Edward I. **Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy**. *Journal of Finance*, v. 23, setembro de 1968.
- ANDERSON, R.L. **The Problem of Autocorrelation in Regression Analysis**. *Journal of the American Statistical Association*, vol.49, n.256, p.113-129, 1954.
- ASTI VERA, Armando. **Metodologia da Pesquisa Científica**. 7ª. Ed. Porto Alegre: Editora Globo, 1983.
- BABBIE, E. **Métodos de Pesquisa de Survey**. Belo Horizonte: Ed. UFMG, 2001.
- BALERA, Wagner. **Sistema de Seguridade Social**. São Paulo: LTr, 2000.
- BOAVENTURA, Edivaldo M. **Metodologia da Pesquisa**. São Paulo: Atlas, 2007.
- BORJESON, L., HOJER, M., DREBORG, K. H., EKVALL, T., FINNVEDEN, G. **“Towards a User’s Guide to Scenarios: a Report on Scenario Type and Scenario Techniques”**, *Environmental Strategies Research*. Stockholm: Royal Institute of Technology, 2005. Disponível em: <[http:// www.infra.kth.se/fms](http://www.infra.kth.se/fms)> Acesso em: 05/02/2008.
- BRASIL. Livro **Branco da Previdência Social**. Ministério da Previdência e Assistência Social. Brasília, 2002. 152. p.
- BRITO, Giovani. A.S.; ASSAF NETO, A. Modelo de Classificação de Risco de Crédito de Empresas. **Revista de Contabilidade e Finanças (RCF – USP)**, São Paulo, v.19, n.46, p.18-29, jan./abr. 2008.

BUARQUE, S. C., **Metodologia e técnicas de construção de cenários globais e regionais**. IPEA, Textos para discussão, fev.2003.

CAPELLETTO, Lucio Rodrigues; MARTINS, Eliseu; CORRAR, Luiz João. **Mensuração do Risco Sistêmico no Setor Bancário com Variáveis Contábeis e Econômicas**. Trabalhos para discussão, n. 169 p. 1-50. Brasília, 2008. Disponível em : [HTTP://www.bcb.gov.br](http://www.bcb.gov.br) . Acessado em: 10/09/2008.

CASCARINO, Richard. **Auditor's Guide to Information Systems Auditing**. Hoboken, New Jersey: Ed. John Wiley & Sons Inc, 2007.

CARVALHO, Alexandre Pinto de; NERI, Marcelo Cortes. Mobilidade Ocupacional e Raça: Origens, Destinos e Riscos dos Afro-Brasileiros. Rio de Janeiro: Fundação Getúlio Vargas, 2000. Disponível em: [HTTP://virtualbib.fgv.br/dspace/handle/123456789/408](http://virtualbib.fgv.br/dspace/handle/123456789/408). Acessado em: 05/07/2008.

CAVUSGIL, S.T; ELVEY-KIRK, L.A. **Mail Server response behavior: A conceptualization motivating factors and empirical study**. European Journal of Marketing, v. 32, n. 11/12, p. 1165-1192, 1998.

CHRISTENSEN, R.; **Log-linear models & Logistic Regression**. New York: Springer-Verlag,1997. 500 p.

COOPER, Donald R.;SCHINDLER, Pamela S. **Métodos de Pesquisa em Administração**. Tradução de Luciana de Oliveira da Rocha. 7ed, Porto Alegre: Bookman, 2003.

CORRAR, Luiz J.; PAULO, Edilson; DIAS FILHO, José Maria. **Análise Multivariada**. São Paulo: Atlas, 2007.

COX, D.R. **Notes on Some Aspects of Regression Analysis**. Journal of the Royal Statistical Society, vol.131, n.3, p.265-279, 1968.

COX, D.R.; HINKLEY, D. V.; **Theoretical statistics**., London: Chapman & Hall, 1986. 174 p.

COX, D. R.; SNELL, E.J. Analysis of binary data. 2. Ed. London: Champman and Hall, 1989.

CRAMER, J. S.; **Logit models from economics and other fields**. Cambridge: Cambridge University, 2003. 184 p.

DATAPREV. Empresa de Tecnologia e Informações da Previdência Social. **Anuário Estatístico Da Previdência Social: Suplemento Histórico – V. 1(1980.1996) – Brasília 1997**.

DATAPREV. **O Modelo Previdenciário Brasileiro**. Rio de Janeiro: 1979.

DATAPREV.**SUIBE – Sistema Unificado de Informações sobre Benefícios** [site intrane WWW-SUIBE]. Rio de Janeiro: março/2007.

DEMO, Pedro. **Introdução à ciência**. São Paulo: Atlas, 1985.

DESLANDES, Suely Ferreira. A CONSTRUÇÃO DO PROJETO DE PESQUISA. In: MINAYO, Cecilia Maria S.; NETO, Otávio C.; GOMES, Romeu. **Pesquisa Social:Teoria, Método e Criatividade**. Petrópolis: Editora Vozes, 1998. 81p.

FRÁGUAS JUNIOR, Renério; ALVES, Tânia. C. T. F. Depressão no Hospital Geral:Estudo de 136 casos. Revista da Associação Médica Brasileira. V.48. n. 3. 2002. p. 225-230.

GIL, Antonio Carlos. **Como elaborar projetos e Pesquisa**. 3ª. Ed. São Paulo: Atlas;1995:58.

GIL, Antônio Carlos. **Métodos e Técnicas de Pesquisa Social**. 5ª.ed. São Paulo: Atlas, 1999.

GIMENO, Suely Godoy A.; SOUZA, J.M.P. Utilização de estratificação e modelo de regressão logística na análise de dados de estudos caso-controle. Revista de Saúde Pública. v.29. n.4.1995.p.283-289.

GOODE, W.J.; HATT P.K. **Métodos em pesquisa Social**. 5ª. Ed. São Paulo: Companhia Editora Nacional,1979.

GUIMARÃES, I.A; NETO, A.C. Reconhecimento de Padrões: Metodologias estatísticas em crédito ao consumidor. Rae-eletrônica. V.1. Número 2, jul-dez/2002. Disponível em: <http://www.rae.com.br/electronica/index.cfm?FuseAction=Artigo&ID=1215&Secao=FINANÇAS2&Volume=1&Numero=2&Ano=2002>.

HAIR, Joseph .F.Jr; ANDERSON, Rolph.E; TATHAM, Ronald. L; BLACK, William. C. **Análise Multivariada de dados**. Tradução Adonai Schlup Sant’Anna e Anselmo Chaves Neto. - 5ª. Ed. - Porto Alegre: Bookman, 2005.

HOSMER, D.W.; LEMESHOW, S.; **Applied logistic regression**. New York: John Wiley, 1989, 307 p.

JANES, J. **On research: Survey construction**. Library Hitech, 1999.

KLEINBAUM, D. G.; **Logistic regression: a self-learning text**. New York: Springer-Verlac,1994. 278 p.

LAKATOS, Eva Maria; MARCONI, Marina de Andrade. **Fundamentos da Metodologia Científica**. 6ª. Ed. São Paulo: Atlas, 2007a.

LAKATOS, Eva Maria; MARCONI, Marina de Andrade. **Técnicas de Pesquisa**. 6ª. Ed. São Paulo: Atlas, 2007b.

LAKATOS, Eva Maria; MARCONI, Marina de Andrade. **Metodologia Científica**. 5ª. Ed. São Paulo: Atlas, 2008.

LAVINAS, L; NICOLL, M. **Atividade e vulnerabilidade: Quais os arranjos familiares em risco?**. Revista de Ciências Sociais. Rio de Janeiro, vol. 49, no.1, 2006, PP. 67-97.

MAKRIDAKIS, SPYROS G., WHEELWRIGHT, S. C., MCGEE, V. E. **Forecasting, Methods and Application**. 2 ed. New York: John Wiley & Sons, 1983.

MAKRIDAKIS, SPYROS,G. **Forecasting, Planning and Strategy for 21st Century**. New York: THE FREE PRESS, 1990. 279p.

MALHOTRA, Naresh. **Pesquisa de Marketing: uma orientação aplicada**. Tradução Laura Bocco. – 4ª ed. – Porto Alegre: Bookman, 2006.

MARQUES, H. R.; MANFROI, J.; CASTILHO, M. A.; NOAL, M. L. **Metodologia da pesquisa e do trabalho científico**. 2. ed. CAMPO GRANDE: EDITORA UCDB, 2006. v. 1. 130 p.

MATIAS-PEREIRA, José. **Manual de Metodologia da Pesquisa Científica**. São Paulo: Atlas, 2007.

McCLAVE, James T.; BENSON, P. George; SINCICH, Terry. **Statistics for Business and Economics**. 8a. ed. Londres: Prentice Hall, 2001.

MENARD, SCOTT. **Applied logistic regression analysis**. 1995.

MICHEL, Maria Helena. **Metodologia e Pesquisa Científica em Ciências Sociais**.São Paulo: Atlas, 2005.

MINAYO, Maria Cecília S.; DESLANDES, Suely F.;NETO, O.C.; GOMES, R. **Pesquisa Social:Teoria, Método e Criatividade**. Petrópolis: Editora Vozes, 1998.

MINUSSI, João Alberto; DAMASCENA, Cláudio; NESS-JUNIOR, Walter Lee. **Um modelo preditivo de Solvência Utilizando Regressão Logística**. In: ENANPAD, 25. 2001, Campinas, Anais. Rio de Janeiro: ANPAD, 2001.

MONTGOMERY,D. C.; PECK, E. A.; **Introduction to linear regression analysis.**, New York: John Wiley, 1992. 527 p.

MORITZ, G. **Planejamento por cenários prospectivos: a construção de um referencial metodológico baseado em casos**. Tese (Doutorado) – UFSC, Florianópolis, 2004.

MPAS. **Dinâmica Socio-econômica e Previdência Sistêmica**. Brasília: 1999.

MPAS. Empresa de Tecnologia e Informações da Previdência Social. **ANUÁRIO ESTATÍSTICO DA PREVIDÊNCIA SOCIAL – Ano 1 (1988.1992)** – Brasília 1993.

MPAS. Empresa de Tecnologia e Informações da Previdência Social. **ANUÁRIO ESTATÍSTICO DA PREVIDÊNCIA SOCIAL – Ano 03**. ANO 2006

MPAS. Empresa de Tecnologia e Informações da Previdência Social. **Boletim Estatístico da Previdência Social - Vol. 12 Nº 07**. Brasília. MPAS, 2006.

MPAS. **Previdência, Assistência Social e Combate à Pobreza**. Brasília, 2000 196 p. (Previdência Social, série Debates) ISBN-85-88219-05-0.

MPAS. **Regulamento da Previdência Social**. Decreto 3048 de 06/05/1999. Brasília 1999, 334. p.

MPAS. **Seguridade e Desenvolvimento: Um projeto para o Brasil**. Brasília:2001.

NAGELKERKE, N. J. D. **A note on a general definition of the coefficient of determination**. *Biometrika*, p. 691-692, 1991.

OHLSON, J. **Financial Ratios and the Probabilistic Prediction of Bankruptcy**. *Journal of Accounting Research*, v. 18, n.1, 1980.

ONUSIC, L.M.; KASSAI, Silvia; VIANA, A. B. N. Comparação dos resultados de utilização de análise por envoltória de dados e regressão logística em modelos de previsão de insolvência: Um estudo aplicado a empresas brasileiras. **Facep Pesquisa – V. 7 – N. 1, 2004**. Disponível em: http://www.facef.br/facefpesquisa/2004/nr1/2_ONUSIC_KASSAI_VIANA.pdf.
Acessado em 11/07/2008.

PANNEL, P.B.W.; PANELL, D.J. **Introducing to social surveying: Pitfalls Potential Problems and Preferred Practices**, 2002.

PEARSON, N.C.S. **Response scales: How many points and what labels?** Disponível em: <<http://datamanagement.scantron.com>>. Acesso em: 20 mar 2008.

PEREIRA, Orlando Mansur; NESS JUNIOR, Walter Lee. **O modelo e-score de previsão de Falências para Empresas de Internet**. Rio de Janeiro: Fundação Getúlio Vargas, 2008. Disponível em: <HTTP://virtualbib.fgv.br/dspace/handle/123456789/1231>. Acessado em: 10/09/2008.

RANCHHOD, A.; ZAROU, F. **Comparing respondents of e-mails and mail surveys: Understanding the implications of technologies**. *Marketing Intelligence & Planning*, v.19, n.4, p. 254-262, 2001.

REZENDE, Solange Oliveira. **Sistemas Inteligentes**. São Paulo: Manole, 2003.

RODRIGUES, Rui Martinho. **Pesquisa Acadêmica**. São Paulo: Atlas, 2007.

ROESH, Sylvia Maria Azevedo. **Projeto de estágio do curso de Administração: guia para pesquisas, projetos, estágios e trabalhos de conclusão de curso**. São Paulo: Atlas, 1996.

SAMPAIO, Rosana. F.; NAVARRO, Albert; MARTIN, M. **Incapacidades laborales: problemas em la reinserción al trabajo**. Cadernos de Saúde Pública. Rio de Janeiro, out-dez, 1999.

SANTOS, Narcisa M. G. dos. **Um estudo prospectivo sobre a Previdência Social Brasileira**. Rio de Janeiro, 2004. Tese – Universidade Federal do Rio de Janeiro, COPPE – 229 p.

SANTOS₁, Alcione M.; SEIXAS, J. M.; PEREIRA, B. B.; MEDRONHO, R. A. **Utilizando Redes Neurais artificiais e Regressão Logística na Predição da Hepatite A**. Revista Brasileira de Epidemiologia. V.8. n.2, p. 117-126, 2008.

SCHWARTZ, P. **A arte da visão de longo prazo: planejando o futuro em um mundo de incertezas**. São Paulo: Best Seller, 2000.

SIMON, Herbert Alexander. **Comportamento Administrativo**. Rio de Janeiro: Fundação Getúlio Vargas, 1957.

SIMON, Herbert Alexander. **A capacidade de decisão e de liderança**. Rio de Janeiro: Editora Fundo de Cultura, 1960.

SINCICH, Terry. **Business Statistics by Example**. 5ed. New Jersey: Prentice-Hall, 1996.

SINPAS. **Reorganização e racionalização da Previdência Social**. Lei no. 6439 de 01/09/1977. Brasília. MPAS. 1977.

STEVENSON, Willian J. **Estatística Aplicada à Administração**. São Paulo: Harbra, 2001.

TABACHNICK, Barbara G.; FIDELL, Linda. S. **Using multivariate statistics**. New York: HarperCollins, 2001.

THOMAS, W.; COOK, R. D.; **Assessing influence on predictions from generalized linear models technometrics**, Alexandria, v.32, p.59-65, 1990.

VENTURA, Magda Maria. **O estudo de Caso como Modalidade de Pesquisa**. Rio de Janeiro: Socerj, 2007.

WANG, HSIAO-FAN. **Intelligent Data Analysis: Developing New Methodologies Through Pattern Discovery and Recovery**. New York: Information Science Reference, 2008.

WANKE, Peter Fernandes. Dinâmica da Estratégia Logística em Empresas Brasileiras. RAE, Rio de Janeiro, v.45, n. 4, p. 22-35, out/dez. 2005.

Yin, R. K. Case Study Research: (1984). Design and Methods. Vol. 5. **Applied Social Research Methods**, ed. Leonard Bickman. Beverly Hills, CA: Sage, 1984.

YIN, Robert K. **Estudo de caso: planejamento e métodos**. 2ª. ed. Porto Alegre: Bookman, 2001.

ZADEH, Lofti. **Outline of a new approach to the analysis of complex systems and decision processes**. IEEE Trans. Syst. Man. Cybern., vol SMC-3 (1973) pp 28-44.

ZEILHOFER, Peter.; OLIVEIRA, Ivani Matos.; KLEMP, S. M.; SANTOS, E.S.; DORES, E. F. G. C. **SIG e regressão logística para mapeamento de risco de contaminação por pesticidas nos mananciais superficiais da bacia do Alto Rio das Mortes-MT**. Anais XIII Simpósio Brasileiro de Sensoriamento Remoto, Florianópolis, Brasil, 21-26 abril 2007, INPE, p. 3623-3630.

8 ANEXOS

8.1 ANEXO A – Auxílio doença Previdenciário

Identificado pelo INSS como “Espécie 31”, é o benefício concedido ao segurado impedido de trabalhar por doença ou acidente por mais de 15 dias consecutivos. No caso dos trabalhadores com carteira assinada, os primeiros 15 dias são pagos pelo empregador, e a Previdência Social paga a partir do 16º dia de afastamento do trabalho. No caso do contribuinte individual (empresário, profissionais liberais, trabalhadores por conta própria, entre outros), a Previdência Social paga todo o período da doença ou do acidente (desde que o trabalhador tenha requerido o benefício).

Para ter direito ao benefício, o trabalhador tem de contribuir para a Previdência Social por, no mínimo, 12 meses. Esse prazo não será exigido em caso de acidente de qualquer natureza (por acidente de trabalho ou fora do trabalho). Para concessão de auxílio-doença é necessário identificar e comprovar a incapacidade através de exame realizado pela perícia médica da Previdência Social.

Terá direito ao benefício sem a necessidade de cumprir o prazo mínimo de contribuição, desde que tenha qualidade de segurado, o trabalhador acometido de tuberculose ativa, hanseníase, alienação mental, neoplasia maligna, cegueira, paralisia irreversível e incapacitante, cardiopatia grave, doença de Parkinson, espondiloartrose anquilosante, nefropatia grave, doença de Paget (osteíte deformante) em estágio avançado, síndrome da deficiência imunológica adquirida (Aids) ou contaminado por radiação (comprovada em laudo médico).

O trabalhador que recebe auxílio-doença é obrigado a realizar exame médico periódico e participar do programa de reabilitação profissional prescrito e custeado pela Previdência Social, sob pena de ter o benefício suspenso.

Não tem direito ao auxílio-doença quem, ao se filiar à Previdência Social, já tiver doença ou lesão que geraria o benefício, a não ser quando a incapacidade resulta do agravamento da enfermidade. Quando o trabalhador perde a qualidade de segurado, as contribuições anteriores só são consideradas para concessão do auxílio-doença após nova filiação à Previdência Social e houver pelo menos quatro contribuições que, somadas às anteriores, totalizem no mínimo 12.

O auxílio-doença deixa de ser pago quando o segurado recupera a capacidade e retorna ao trabalho ou quando o benefício se transforma em aposentadoria por invalidez.

Os critérios para a concessão do benefício foram obtidos de http://www.inss.gov.br/pg_secundarias/beneficios_06_01.asp.

8.2 ANEXO B – Aposentadoria por invalidez acidente do trabalho

Identificado pelo INSS como “Espécie 32” ou “Espécie 92”, é o benefício concedido aos trabalhadores que, por doença ou acidente, forem considerados pela perícia médica da Previdência Social incapacitados para exercer suas atividades ou outro tipo de serviço que lhes garanta o sustento.

Não tem direito à aposentadoria por invalidez quem, ao se filiar à Previdência Social, já tiver doença ou lesão que geraria o benefício, a não ser quando a incapacidade resultar no agravamento da enfermidade.

Quem recebe aposentadoria por invalidez é obrigado a passar por perícia médica de dois em dois anos, se não, o benefício será suspenso. A aposentadoria deixa de ser paga quando o segurado recupera a capacidade e volta ao trabalho.

Para ter direito ao benefício, o trabalhador tem que contribuir para a Previdência Social por no mínimo 12 meses, no caso de doença. Se for acidente, esse prazo de carência não é exigido, mas é preciso estar inscrito na Previdência Social (http://www.inss.gov.br/pg_secundarias/beneficios_06.asp).

8.3 ANEXO C - Auxílio por Acidente de Trabalho

Identificado pelo INSS como “Espécie 91”, é o benefício concedido ao segurado incapacitado para o trabalho em decorrência de acidente de trabalho ou de doença profissional. Considera-se acidente de trabalho aquele ocorrido no exercício de atividades profissionais a serviço da empresa (típico) ou ocorrido no trajeto casa-trabalho-casa (de trajeto).

Têm direito ao auxílio-doença acidentário o empregado, o trabalhador avulso, o médico-residente e o segurado especial. A concessão do auxílio-doença acidentário não exige tempo mínimo de contribuição. Ao trabalhador que recebe auxílio-doença, a Previdência oferece o programa de reabilitação profissional.

A comunicação de acidente de trabalho ou doença profissional será feita à Previdência Social em formulário próprio denominado CAT (Comunicação de Acidente de Trabalho) e deverá ser emitida pela empresa ou pelo próprio trabalhador, por seus dependentes, pela entidade sindical, pelo médico ou por autoridade (magistrados, membros do Ministério Público e dos serviços jurídicos da União, dos estados e do Distrito Federal e comandantes de unidades do Exército, da Marinha, da Aeronáutica, do Corpo de Bombeiros e da Polícia Militar). O formulário preenchido tem que ser entregue em uma Agência da Previdência Social pelo emitente.

A retomada de tratamento e o afastamento por agravamento de lesão decorrentes de acidente de trabalho ou doença profissional têm de ser comunicados à Previdência Social em formulário próprio. Nessa CAT deverão constar as

informações da época do acidente e os dados atualizados do novo afastamento (último dia trabalhado, atestado médico e data da emissão).

Também devem ser informadas à Previdência Social por meio da CAT mortes de segurados decorrentes de acidente de trabalho ou doença ocupacional.

A empresa é obrigada a informar à Previdência Social acidentes de trabalho ocorridos com seus funcionários, mesmo que não haja afastamento das atividades, até o primeiro dia útil seguinte ao da ocorrência. Em caso de morte, a comunicação deve ser imediata. A empresa que não informar acidentes de trabalho está sujeita à multa. Nos primeiros 15 dias de afastamento, o salário do trabalhador é pago pela empresa, depois, a Previdência Social é responsável pelo pagamento. Enquanto recebe auxílio-doença por acidente de trabalho ou doença ocupacional, o trabalhador é considerado licenciado e terá estabilidade por 12 meses após o retorno às atividades.

O auxílio-doença deixa de ser pago quando o segurado recupera a capacidade e retorna ao trabalho ou quando o benefício se transforma em aposentadoria por invalidez.

9 APÊNDICES

9.1 APÊNDICE A - Questionário

Prezado respondente, o questionário a seguir destina-se a complementar uma pesquisa, inicialmente de caráter acadêmico, cujo resultado poderá ser utilizado nos processos de Gestão e Auditoria da Previdência Social, especificamente na identificação de padrões de irregularidade nos processos de Concessão e Manutenção dos Benefícios por incapacidade (Auxílio-acidente do Trabalho e Aposentadoria por Invalidez), possibilitando a identificação de critérios específicos dos benefícios por incapacidade, os quais permitirão agilidade nas ações de verificação e acompanhamento sobre benefícios específicos, oferecendo alternativas que poderão contribuir para a redução financeira na área acidentária do custeio da Previdência Social.

Estas informações integrarão a dissertação de Mestrado em Administração e Desenvolvimento Empresarial conduzida pelo funcionário Jonias dos Santos Bueno (Jonias.Bueno@Previdencia.gov.br), integrante da Dataprev, lotado no Rio de Janeiro, com o apoio dos Órgãos de Gestão do MPAS e Dataprev.

A dissertação intitulada “Indicadores Antecedentes na Área de Acidentes de Trabalho na Previdência Social Brasileira” utilizará as informações prestadas para construir uma base de conceitos e fundamentar o trabalho acadêmico, já que não foram localizados referenciais acadêmicos sobre o assunto, nesta situação, a ciência adota a opinião de especialistas como embasamento inicial na condução das pesquisas científicas.

Desde já agradeço a sua valiosa contribuição e coloco-me à disposição para possíveis esclarecimentos adicionais.

Relação dos fatores que serão avaliados no estudo:

Fator	
1	Idade do Segurado na concessão do Benefício
2	Idade do Segurado em Janeiro/ 2008
3	Sexo do Segurado
4	Tempo de Contribuição
5	Espécie do Benefício
6	Forma de Filiação ao Regime Geral de Previdência Social (RGPS)
7	Atividade desempenhada pelo segurado (CBO)
8	Quantidade de dependentes para Imposto de Renda (IR)
9	Renda Mensal Inicial prevista em Salários-Mínimos
10	Tipo de acidente
11	Ramo de atividade
12	Quantidade de revisões da Renda Mensal Inicial
13	Se tem Perícia Médica realizada
14	Unidade da Federação onde o benefício foi concedido
15	Unidade da Federação onde o segurado trabalhou
16	Unidade da Federação onde o benefício é recebido
17	Tipo de Sistema de Concessão
18	Uso de informações do CNIS
19	Tipo de despacho do benefício
20	Tempo de Recebimento do Benefício

Este estudo levará em consideração todos os fatores em conjunto, buscando a influência de cada fator na identificação de um **benefício por incapacidade** como Ativo ou Cessado, permitindo também avaliar a conformidade dos processos de concessão e manutenção dos benefícios selecionados para o estudo, para tanto, necessitamos de sua opinião quanto à utilização ou não dos fatores citados acima no âmbito do estudo.

I) Assinale com um "X" a opção que melhor representa a sua opinião sobre a questão formulada.

1. A **idade** do beneficiário na época do requerimento do Benefício por incapacidade deve ser considerada no estudo.

Concordo Totalmente	Concordo Muito	Concordo	Neutro	Discordo	Discordo Muito	Discordo Totalmente
1	2	3	4	5	6	7

2. A **idade** do beneficiário em Janeiro/2008 (período final do estudo) deve ser considerada no estudo.

Concordo Totalmente	Concordo Muito	Concordo	Neutro	Discordo	Discordo Muito	Discordo Totalmente
1	2	3	4	5	6	7

3. O **sexo** do beneficiário de um Benefício por incapacidade deve ser considerado no estudo.

Concordo Totalmente	Concordo Muito	Concordo	Neutro	Discordo	Discordo Muito	Discordo Totalmente
1	2	3	4	5	6	7

4. O **tempo de contribuição** do segurado em um Benefício Acidentário deve ser considerado no estudo.

Concordo Totalmente	Concordo Muito	Concordo	Neutro	Discordo	Discordo Muito	Discordo Totalmente
1	2	3	4	5	6	7

5. A **espécie** de um Benefício Acidentário (espécies 91 e 92) deve ser considerada no estudo.

Concordo Totalmente	Concordo Muito	Concordo	Neutro	Discordo	Discordo Muito	Discordo Totalmente
1	2	3	4	5	6	7

6. A **forma de filiação** do segurado ao Sistema Previdenciário deve ser considerada no estudo.

Concordo Totalmente	Concordo Muito	Concordo	Neutro	Discordo	Discordo Muito	Discordo Totalmente
1	2	3	4	5	6	7

7. A **atividade profissional desempenhada** (Classificação Brasileira de Ocupações – CBO) pelo Beneficiário deve ser considerada no estudo.

Concordo Totalmente	Concordo Muito	Concordo	Neutro	Discordo	Discordo Muito	Discordo Totalmente
1	2	3	4	5	6	7

8. A **quantidade de dependentes** válidos para Imposto de Renda informada pelo Beneficiário deve ser considerada no estudo.

Concordo Totalmente	Concordo Muito	Concordo	Neutro	Discordo	Discordo Muito	Discordo Totalmente
1	2	3	4	5	6	7

9. A **Renda Mensal Inicial**, em salários mínimos, deve ser considerada no estudo.

Concordo Totalmente	Concordo Muito	Concordo	Neutro	Discordo	Discordo Muito	Discordo Totalmente
1	2	3	4	5	6	7

10. O **tipo de acidente informado** (espécies 91 e 92) deve ser considerado no estudo.

Concordo Totalmente	Concordo Muito	Concordo	Neutro	Discordo	Discordo Muito	Discordo Totalmente
1	2	3	4	5	6	7

- 11. O Ramo de Atividade** dos benefícios por incapacidade solicitados ou concedidos para um Trabalhador deve ser considerado no estudo.

Concordo Totalmente	Concordo Muito	Concordo	Neutro	Discordo	Discordo Muito	Discordo Totalmente
1	2	3	4	5	6	7

- 12. O Quantidade de revisões da Renda Mensal Inicial** sofridas por um benefício Acidentário deve ser considerada no estudo.

Concordo Totalmente	Concordo Muito	Concordo	Neutro	Discordo	Discordo Muito	Discordo Totalmente
1	2	3	4	5	6	7

- 13. A realização de Perícia Médica** em benefícios por incapacidade deve ser considerada no estudo.

Concordo Totalmente	Concordo Muito	Concordo	Neutro	Discordo	Discordo Muito	Discordo Totalmente
1	2	3	4	5	6	7

- 14. A Unidade da Federação (UF)** na qual **foi concedido** um Benefício por incapacidade deve ser considerada no estudo.

Concordo Totalmente	Concordo Muito	Concordo	Neutro	Discordo	Discordo Muito	Discordo Totalmente
1	2	3	4	5	6	7

- 15. A Unidade da Federação** na qual **o Segurado exerceu atividade profissional** deve ser considerada no estudo.

Concordo Totalmente	Concordo Muito	Concordo	Neutro	Discordo	Discordo Muito	Discordo Totalmente
1	2	3	4	5	6	7

- 16. A Unidade da Federação** onde o Benefício **é mantido pela Previdência Social** deve ser considerada no estudo.

Concordo Totalmente	Concordo Muito	Concordo	Neutro	Discordo	Discordo Muito	Discordo Totalmente
1	2	3	4	5	6	7

- 17. O Tipo de Sistema de Concessão** atribuído ao benefício deve ser considerado no estudo.

Concordo Totalmente	Concordo Muito	Concordo	Neutro	Discordo	Discordo Muito	Discordo Totalmente
1	2	3	4	5	6	7

- 18. A utilização das informações do Cadastro Nacional de Informações Sociais (CNIS)** na concessão dos benefícios por incapacidade deve ser considerada no estudo.

Concordo Totalmente	Concordo Muito	Concordo	Neutro	Discordo	Discordo Muito	Discordo Totalmente
1	2	3	4	5	6	7

19. O Tipo de Despacho comandado para a concessão de benefícios Acidentários deve ser considerado no estudo.

Concordo Totalmente	Concordo Muito	Concordo	Neutro	Discordo	Discordo Muito	Discordo Totalmente
1	2	3	4	5	6	7

20. O Tempo de Recebimento do benefício Acidentário deve ser considerado no estudo.

Concordo Totalmente	Concordo Muito	Concordo	Neutro	Discordo	Discordo Muito	Discordo Totalmente
1	2	3	4	5	6	7

II) Nesta seção, solicitamos que assinale com um “X” a opção relativa a seus dados pessoais.

1) Tempo de atuação na área de Benefícios por Incapacidade na Previdência Social:

Menos de 05 anos	De 05 a 10 anos	De 10 a 15 anos	De 15 a 20 anos	20 anos ou mais
1	2	3	4	5

2) Nível de atuação sobre os Benefícios por Incapacidade na Previdência Social:
(Considerar o nível em que o respondente atuou durante o maior tempo)

Gestão de Tecnologia da Informação	Gestão de Negócios Previdenciários	Execução em Tecnologia da Informação	Execução em Negócios Previdenciários
1	2	3	4

3) Área de atuação na Previdência Social:
(Considerar a área em que o respondente atuou durante o maior tempo)

Administração Central MPAS	Agência da Previdência Social	Administração Central Dataprev	Regionais Dataprev
1	2	3	4

9.2 APÊNDICE B – Variáveis categóricas do questionário

Questão 3 - O **sexo** do beneficiário de um Benefício por incapacidade deve ser considerado no estudo.

Sexo	
Validade	Descrição
1	MASCULINO
3	FEMININO
9	NAO INFORMADO

Questão 5 - A **espécie** de um Benefício Acidentário (espécies 91 e 92) deve ser considerada no estudo

Espécies de Benefícios	
Validade	Descrição
1	PENSAO POR MORTE DE TRABALHADOR RURAL
2	PENSAO POR MORTE ACIDENTARIA-TRAB. RURAL
3	PENSAO POR MORTE DE EMPREGADOR RURAL
4	APOSENTADORIA POR INVALIDEZ-TRAB. RURAL
5	APOSENT. INVALIDEZ ACIDENTARIA-TRAB.RUR.
6	APOSENT. INVALIDEZ EMPREGADOR RURAL
7	APOSENTADORIA POR VELHICE - TRAB. RURAL
8	APOSENT. POR IDADE - EMPREGADOR RURAL
9	COMPL. ACIDENTE TRABALHO P/TRAB. (RURAL)
10	AUXILIO DOENCA ACIDENTARIO - TRAB. RURAL
11	AMPARO PREVIDENC. INVALIDEZ- TRAB. RURAL
12	AMPARO PREVIDENC. IDADE - TRAB. RURAL
13	AUXILIO DOENCA - TRABALHADOR RURAL
15	AUXILIO RECLUSAO - TRABALHADOR RURAL
19	PENSAO DE ESTUDANTE (LEI 7.004/82)
20	PENSAO POR MORTE DE EX-DIPLOMATA
21	PENSAO POR MORTE PREVIDENCIARIA
22	PENSAO POR MORTE ESTATUTARIA
23	PENSAO POR MORTE DE EX-COMBATENTE
24	PENSAO ESPECIAL (ATO INSTITUCIONAL)
25	AUXILIO RECLUSAO
26	PENSAO POR MORTE ESPECIAL
27	PENSAO MORTE SERVIDOR PUBLICO FEDERAL
28	PENSAO POR MORTE REGIME GERAL
29	PENSAO POR MORTE EX-COMBATENTE MARITIMO
30	RENDA MENSAL VITALICIA POR INCAPACIDADE
31	AUXILIO DOENCA PREVIDENCIARIO
32	APOSENTADORIA INVALIDEZ PREVIDENCIARIA
33	APOSENTADORIA INVALIDEZ AERONAUTA
34	APOSENT. INVAL. EX-COMBATENTE MARITIMO
35	AUXILIO-DOENCA DO EX-COMBATENTE
36	AUXILIO ACIDENTE PREVIDENCIARIO
37	APOSENTADORIA EXTRANUMERARIO CAPIN
38	APOSENT. EXTRANUM. FUNCIONARIO PUBLICO
39	AUXILIO INVALIDEZ ESTUDANTE
40	RENDA MENSAL VITALICIA POR IDADE
41	APOSENTADORIA POR IDADE

42	APOSENTADORIA POR TEMPO DE CONTRIBUICAO
43	APOSENT. POR TEMPO SERVICO EX-COMBATENTE
44	APOSENTADORIA ESPECIAL DE AERONAUTA
45	APOSENT. TEMPO SERVICO JORNALISTA
46	APOSENTADORIA ESPECIAL
47	ABONO PERMANENCIA EM SERVICO - 35 ANOS
48	ABONO PERMANENCIA EM SERVICO - 30 ANOS
49	APOSENTADORIA ORDINARIA
50	AUXILIO DOENCA EXTINTO PLANO BASICO
51	APOSENT. INVALIDEZ EXTINTO PLANO BASICO
52	APOSENT. IDADE EXTINTO PLANO BASICO
53	AUXILIO RECLUSAO EXTINTO PLANO BASICO
54	PENSAO ESPECIAL VITALICIA - LEI 9793/99
55	PENSAO POR MORTE EXTINTO PLANO BASICO
56	PENSAO VITALICIA SINDROME TALIDOMIDA
57	APOSENT. TEMPO DE SERVICO DE PROFESSOR
58	APOSENTADORIA DE ANISTIADOS
59	PENSAO POR MORTE DE ANISTIADOS
60	PENSAO ESPECIAL PORTADOR DE SIDA
61	AUXILIO NATALIDADE
62	AUXILIO FUNERAL
63	AUXILIO-FUNERAL TRABALHADOR RURAL
64	AUXILIO-FUNERAL EMPREGADOR RURAL
65	PECULIO ESPECIAL SERVIDOR AUTARQUICO
66	PEC. ESP. SERVIDOR AUTARQUICO
67	PECULIO OBRIGATORIO EX-IPASE
68	PECULIO ESPECIAL DE APOSENTADOS
69	PECULIO DE ESTUDANTE
70	RESTITUICAO CONTRIB. P/SEG. S/CARENCIA
71	SALARIO-FAMILIA PREVIDENCIARIO
72	APOSENT. TEMPO SERVICO - LEI DE GUERRA
73	SALARIO FAMILIA ESTATUTARIO
74	COMPLEMENTO DE PENSAO A CONTA DA UNIAO
75	COMPLEMENTO DE APOSENT. A CONTA DA UNIAO
76	SALARIO FAMILIA ESTATUTARIO
77	SALARIO FAM. ESTATUTARIO SERVIDOR SINPAS
78	APOSENTADORIA IDADE - LEI DE GUERRA
79	VANTAGENS DE SERVIDOR APOSENTADO
80	AUXILIO SALARIO MATERNIDADE
81	APOSENTADORIA COMPULSORIA EX-SASSE
82	APOSENTADORIA TEMPO DE SERVICO EX-SASSE
83	APOSENTADORIA POR INVALIDEZ EX-SASSE
84	PENSAO POR MORTE EX-SASSE
85	PENSAO VITALICIA SERINGUEIROS
86	PENSAO VITALICIA DEPENDENTES SERINGUEIRO
87	AMP. SOCIAL PESSOA PORTADORA DEFICIENCIA
88	AMPARO SOCIAL AO IDOSO
89	PENSAO ESP. VITIMAS HEMODIALISE-CARUARU
90	SIMPLES ASSIST. MEDICA P/ ACIDENTE TRAB.
91	AUXILIO DOENCA POR ACIDENTE DO TRABALHO
92	APOSENT. INVALIDEZ ACIDENTE TRABALHO
93	PENSAO POR MORTE ACIDENTE DO TRABALHO
94	AUXILIO ACIDENTE
95	AUXILIO SUPLEMENTAR ACIDENTE TRABALHO
96	PENSAO ESPECIAL-MEDIDA PROVISORIA 373/07
97	PECULIO POR MORTE ACIDENTE DO TRABALHO
98	ABONO ANUAL DE ACIDENTE DE TRABALHO
99	AFASTAMENTO ATE 15 DIAS ACIDENTE TRAB.

Questão 6 - A **forma de filiação** do segurado ao Sistema Previdenciário deve ser considerada no estudo

Formas de filiação à Previdência Social	
Validade	Descrição
1	EMPREGADO
2	TRABALHADOR AVULSO
3	EMPRESARIO
4	DOMESTICO
5	FACULTATIVO
6	EQUIPARADO A AUTONOMO
7	SEGURADO ESPECIAL
8	AUTONOMO
9	OPTANTE PELA LEI 6.184/74
0	DESEMPREGADO

Questão 10 - O **tipo de acidente informado** (espécies 91 e 92) deve ser considerado no estudo

Tipos de Acidentes de trabalho	
Validade	Descrição
1	ACIDENTE TIPICO NO LOCAL DE TRABALHO
2	DOENCA PROFISSIONAL OU DO TRABALHO
3	ACIDENTE NO TRAJETO

Questão 11 - O **Ramo de Atividade** dos benefícios por incapacidade solicitados ou concedidos para um Trabalhador deve ser considerado no estudo.

Ramos de atividade dos Trabalhadores	
Validade	Descrição
1	BANCARIO
2	COMERCARIO
3	TRANSPORTES E CARGA
4	FERROVIARIO
5	INDUSTRIARIO
6	MARITIMO
7	SERVIDOR PUBLICO
8	RURAL
9	IRRELEVANTE

Questão 17 - O **Tipo de Sistema de Concessão** atribuído ao benefício deve ser considerado no estudo.

Sistemas de Concessão	
Validade	Descrição
1	CONCESSAO FORMULARIO CCE
2	CONCESSAO PRISMA
3	CONCESSAO FORMULARIO CB/RSC
4	CONCESSAO PRISMA/SUB
5	CONCESSAO ON-LINE
6	CONCESSAO DESCENTRALIZADA, SEM INDICE
7	CONCESSAO DESCENTRALIZADA
8	CONCESSAO ON-LINE, SEM INDICE
9	PRISMA/SUB REGIONAL (ASSOCIADOS)
10	INCLUSAO BENEFICIO FORA DE CADASTRO
11	CONCESSAO PELA INTERNET

Questão 19 - O **Tipo de Despacho** comandado para a concessão de benefícios Acidentários deve ser considerado no estudo.

Tipos de despachos aceitos pela Previdência Social	
Validade	Descrição
0	CONCESSAO NORMAL
1	CONCESSAO C/JUSTIFICACAO ADMINISTRATIVA
2	CONCESSAO COM DILIGENCIA (RD OU SP)
3	CONCESSAO EM FASE RECURSAL
4	CONCESSAO DECORRENTE DE ACAO JUDICIAL
5	CONC. DECORRENTE REVISAO ADMINISTRATIVA
6	REABERTURA DE PROCESSO ENCERRADO
7	CONC. PENSAO ESTAT. COM PECULIO ESPECIAL
8	CONC. COM BASE NO ARTIGO 180 DO RBPS
9	CONC. BASE ARTIGO 27 INCISO II DO RBPS
10	CONCESSAO COM CONVERSAO TEMPO DE SERVICO
11	DESPACHO RESPONSABILIDADE REGIONAIS INSS
12	DESPACHO DISPENSA DE CPF E 25% DESCONTO
13	CONC. AUXILIO-DOENCA COM OUTRA ATIVIDADE
14	CONCESSAO COM MULTIPLA ATIVIDADE
15	CONC. COM BASE NO ARTIGO 183 DO RBPS
16	ACIDENTE QUALQUER NATUREZA P/ AUX-DOENCA
17	CONC. COM BASE ARTIGO 35 DA LEI 8213/91
18	CONCESSAO BASE ARTIGO 122 M.P.1523/97
19	CONC. S/VERIFICACAO DA PERDA QUALIDADE
20	CONCLUSAO MEDICA CONTRARIA
21	INSCRICAO INDEVIDA
22	PERDA DE QUALIDADE DE SEGURADO
23	FALTA DE PERIODO DE CARENCIA
24	FALTA DE QUALIDADE DE DEPENDENTE
25	FALTA DE TEMPO DE SERVICO
26	CAUSAS DIVERSAS
27	FALTA DE VINCULO EMPREGATORIO
28	FALTA NEXO ENTRE ACIDENTE E TRABALHO
29	FALTA NEXO ACIDENTE DE TRABALHO/TRAJETO
30	FALTA NEXO LESAO ACIDENTE/DOENCA TRAB.
31	NAO ATINGIDA IDADE MINIMA APOSENTADORIA
35	INDEFERIMENTO ON-LINE
40	NAO COMPARECIMENTO DO REQUERENTE

41	NAO CUMPRIMENTO DE EXIGENCIAS
42	DESISTENCIA DO REQUERENTE
43	RESTABELECIMENTO DO BENEFICIO ANTERIOR
44	ALTA DE TRATAMENTO ATE 15 DIAS
45	SIMPLES ASSISTENCIA MEDICA
46	MORTE POR ACIDENTE DE TRABALHO
50	TRANSFERENCIA DE BENEFICIO EM CONCESSAO
60	CANCELAMENTO DE PECULIO
99	SIMPLES CONTAGEM DE TEMPO DE SERVICO

9.3 APÊNDICE C – Processamento dos questionários

		Questao_1			
		Freqüência	Percentual	Percentual válido	Percentual acumulado
Válidos	NEUTRO	11	8,7	8,7	8,7
	CONCORDO	23	18,1	18,1	26,8
	CONCORDO MUITO	27	21,3	21,3	48,0
	CONCORDO TOTALMENTE	66	52,0	52,0	100,0
	Total	127	100,0	100,0	

		Questao_2			
		Freqüência	Percentual	Percentual válido	Percentual acumulado
Válidos	DISCORDO MUITO	3	2,4	2,4	2,4
	DISCORDO	24	18,9	18,9	21,3
	NEUTRO	65	51,2	51,2	72,4
	CONCORDO	17	13,4	13,4	85,8
	CONCORDO MUITO	8	6,3	6,3	92,1
	CONCORDO TOTALMENTE	10	7,9	7,9	100,0
	Total	127	100,0	100,0	

		Questao_3			
		Freqüência	Percentual	Percentual válido	Percentual acumulado
Válidos	DISCORDO TOTALMENTE	2	1,6	1,6	1,6
	DISCORDO	1	,8	,8	2,4
	NEUTRO	2	1,6	1,6	3,9
	CONCORDO	5	3,9	3,9	7,9
	CONCORDO MUITO	26	20,5	20,5	28,3
	CONCORDO TOTALMENTE	91	71,7	71,7	100,0
	Total	127	100,0	100,0	

		Questao_4			
		Freqüência	Percentual	Percentual válido	Percentual acumulado
Válidos	NEUTRO	3	2,4	2,4	2,4
	CONCORDO	10	7,9	7,9	10,2
	CONCORDO MUITO	41	32,3	32,3	42,5
	CONCORDO TOTALMENTE	73	57,5	57,5	100,0
	Total	127	100,0	100,0	

		Questao_5			
		Freqüência	Percentual	Percentual válido	Percentual acumulado
Válidos	NEUTRO	2	1,6	1,6	1,6
	CONCORDO	1	,8	,8	2,4
	CONCORDO MUITO	64	50,4	50,4	52,8
	CONCORDO TOTALMENTE	60	47,2	47,2	100,0
	Total	127	100,0	100,0	

Questao_6					
		Freqüência	Percentual	Percentual válido	Percentual acumulado
Válidos	DISCORDO TOTALMENTE	2	1,6	1,6	1,6
	NEUTRO	1	,8	,8	2,4
	CONCORDO	7	5,5	5,5	7,9
	CONCORDO MUITO	72	56,7	56,7	64,6
	CONCORDO TOTALMENTE	45	35,4	35,4	100,0
	Total	127	100,0	100,0	

Questao_7					
		Freqüência	Percentual	Percentual válido	Percentual acumulado
Válidos	DISCORDO	2	1,6	1,6	1,6
	NEUTRO	32	25,2	25,2	26,8
	CONCORDO	30	23,6	23,6	50,4
	CONCORDO MUITO	26	20,5	20,5	70,9
	CONCORDO TOTALMENTE	37	29,1	29,1	100,0
	Total	127	100,0	100,0	

Questao_8					
		Freqüência	Percentual	Percentual válido	Percentual acumulado
Válidos	DISCORDO	15	11,8	11,8	11,8
	NEUTRO	47	37,0	37,0	48,8
	CONCORDO	42	33,1	33,1	81,9
	CONCORDO MUITO	13	10,2	10,2	92,1
	CONCORDO TOTALMENTE	10	7,9	7,9	100,0
	Total	127	100,0	100,0	

Questao_9					
		Freqüência	Percentual	Percentual válido	Percentual acumulado
Válidos	NEUTRO	7	5,5	5,5	5,5
	CONCORDO	23	18,1	18,1	23,6
	CONCORDO MUITO	35	27,6	27,6	51,2
	CONCORDO TOTALMENTE	62	48,8	48,8	100,0
	Total	127	100,0	100,0	

Questao_10					
		Freqüência	Percentual	Percentual válido	Percentual acumulado
Válidos	NEUTRO	7	5,5	5,5	5,5
	CONCORDO	9	7,1	7,1	12,6
	CONCORDO MUITO	34	26,8	26,8	39,4
	CONCORDO TOTALMENTE	77	60,6	60,6	100,0
	Total	127	100,0	100,0	

Questao_11					
		Freqüência	Percentual	Percentual válido	Percentual acumulado
Válidos	DISCORDO	9	7,1	7,1	7,1
	NEUTRO	41	32,3	32,3	39,4
	CONCORDO	44	34,6	34,6	74,0
	CONCORDO MUITO	19	15,0	15,0	89,0
	CONCORDO TOTALMENTE	14	11,0	11,0	100,0
	Total	127	100,0	100,0	

Questao_12					
		Freqüência	Percentual	Percentual válido	Percentual acumulado
Válidos	DISCORDO	4	3,1	3,1	3,1
	NEUTRO	53	41,7	41,7	44,9
	CONCORDO	36	28,3	28,3	73,2
	CONCORDO MUITO	13	10,2	10,2	83,5
	CONCORDO TOTALMENTE	21	16,5	16,5	100,0
	Total	127	100,0	100,0	

Questao_13					
		Freqüência	Percentual	Percentual válido	Percentual acumulado
Válidos	CONCORDO	12	9,4	9,4	9,4
	CONCORDO MUITO	16	12,6	12,6	22,0
	CONCORDO TOTALMENTE	99	78,0	78,0	100,0
	Total	127	100,0	100,0	

Questao_14					
		Freqüência	Percentual	Percentual válido	Percentual acumulado
Válidos	NEUTRO	35	27,6	27,6	27,6
	CONCORDO	17	13,4	13,4	40,9
	CONCORDO MUITO	23	18,1	18,1	59,1
	CONCORDO TOTALMENTE	52	40,9	40,9	100,0
	Total	127	100,0	100,0	

Questao_15					
		Freqüência	Percentual	Percentual válido	Percentual acumulado
Válidos	NEUTRO	2	1,6	1,6	1,6
	CONCORDO	19	15,0	15,0	16,5
	CONCORDO MUITO	48	37,8	37,8	54,3
	CONCORDO TOTALMENTE	58	45,7	45,7	100,0
	Total	127	100,0	100,0	

		<u>Questao_16</u>			
		Freqüência	Percentual	Percentual válido	Percentual acumulado
Válidos	NEUTRO	29	22,8	22,8	22,8
	CONCORDO	31	24,4	24,4	47,2
	CONCORDO MUITO	47	37,0	37,0	84,3
	CONCORDO TOTALMENTE	20	15,7	15,7	100,0
	Total	127	100,0	100,0	100,0

		<u>Questao_17</u>			
		Freqüência	Percentual	Percentual válido	Percentual acumulado
Válidos	NEUTRO	37	29,1	29,1	29,1
	CONCORDO	28	22,0	22,0	51,2
	CONCORDO MUITO	46	36,2	36,2	87,4
	CONCORDO TOTALMENTE	16	12,6	12,6	100,0
	Total	127	100,0	100,0	100,0

		<u>Questao_18</u>			
		Freqüência	Percentual	Percentual válido	Percentual acumulado
Válidos	CONCORDO	2	1,6	1,6	1,6
	CONCORDO MUITO	15	11,8	11,8	13,4
	CONCORDO TOTALMENTE	110	86,6	86,6	100,0
	Total	127	100,0	100,0	100,0

		<u>Questao_19</u>			
		Freqüência	Percentual	Percentual válido	Percentual acumulado
Válidos	CONCORDO	3	2,4	2,4	2,4
	CONCORDO MUITO	27	21,3	21,3	23,6
	CONCORDO TOTALMENTE	97	76,4	76,4	100,0
	Total	127	100,0	100,0	100,0

		<u>Questao_20</u>			
		Freqüência	Percentual	Percentual válido	Percentual acumulado
Válidos	DISCORDO TOTALMENTE	1	0,8	0,8	0,8
	DISCORDO	2	1,6	1,6	2,4
	CONCORDO	4	3,1	3,1	5,5
	CONCORDO MUITO	29	22,8	22,8	28,3
	CONCORDO TOTALMENTE	91	71,7	71,7	100,0
	Total	127	100,0	100,0	100,0

		Demogr_1			
		Frequência	Percentual	Percentual válido	Percentual acumulado
Válidos	Menos de 5 anos	4	3,1	3,1	3,1
	de 5 a 10 anos	8	6,3	6,3	9,4
	de 10 a 15 anos	22	17,3	17,3	26,8
	de 15 a 20 anos	33	26,0	26,0	52,8
	mais de 20 anos	60	47,2	47,2	100,0
	Total	127	100,0	100,0	

		Demogr_2			
		Frequência	Percentual	Percentual válido	Percentual acumulado
Válidos	Gestao de TI	8	6,3	6,3	6,3
	Gestao INSS	64	50,4	50,4	56,7
	Execucao TI	8	6,3	6,3	63,0
	Execucao INSS	47	37,0	37,0	100,0
	Total	127	100,0	100,0	

		Demogr_3			
		Frequência	Percentual	Percentual válido	Percentual acumulado
Válidos	MPAS CENTRAL	73	57,5	57,5	57,5
	APS	39	30,7	30,7	88,2
	DATAPREV CENTRAL	8	6,3	6,3	94,5
	DATAPREV REGIONAL	7	5,5	5,5	100,0
	Total	127	100,0	100,0	

9.4 APÊNDICE D – Variáveis categóricas agrupadas

Codificação das variáveis agrupadas

		Frequência				
		(1)	(2)	(3)	(4)	(5)
CS_SISTEMA_CONCES_GRUPO	4	492	1,000	,000	,000	
	5	694	,000	1,000	,000	
	7	134	,000	,000	1,000	
	8	218	,000	,000	,000	
	99	54	,000	,000	,000	
UnidadeTerritorial	1	101	1,000	,000	,000	
	2	305	,000	1,000	,000	
	3	968	,000	,000	1,000	
	4	149	,000	,000	,000	
	5	69	,000	,000	,000	
CS_FORMA_FILIACAO_GRUPO	0	655	1,000	,000	,000	
	1	506	,000	1,000	,000	
	7	84	,000	,000	1,000	
	8	260	,000	,000	,000	
	99	87	,000	,000	,000	
CS_ESPECIE	31	823	1,000	,000	,000	
	32	685	,000	1,000	,000	
	91	53	,000	,000	1,000	
	92	31	,000	,000	,000	
CS_RAMO_ATIVIDADE_GRUPO	2	1081	1,000	,000	,000	
	5	286	,000	1,000	,000	
	8	109	,000	,000	1,000	
	99	116	,000	,000	,000	
CS_MOT_DESP_GRUPO	0	1574	1,000			
	99	18	,000			
CS_PA	0	1573	1,000			
	1	19	,000			
CS_ACIDENTE_GRUPO	0	1511	1,000			
	99	81	,000			
IN_PERICIA_MEDICA	0	19	1,000			
	1	1573	,000			
CS_TIPO_BENEF_GRUPO	0	1580	1,000			
	99	12	,000			
CS_BATIMENTO_CNIS_GRUPO	0	1336	1,000			
	99	256	,000			
CS_SEXO	1	1033	1,000			
	3	559	,000			

9.5 APÊNDICE E – Análise dos resultados entre 50 e 100%

- Distribuição dos Benefícios por Espécie (CS_ESPECIE)

Especie	Descrição	Quantidade	Valor gerado (R\$)	% do Valor gerado
31	AUXILIO DOENCA PREVIDENCIARIO	53.070	958.028.602,87	71,75%
32	APOSENTADORIA INVALIDEZ PREVIDENCIARIA	10.726	299.110.452,61	22,40%
91	AUXILIO DOENCA POR ACIDENTE DO TRABALHO	804	25.031.575,66	1,87%
92	APOSENT. INVALIDEZ ACIDENTE TRABALHO	1.184	52.982.047,44	3,97%
Total		65.784	1.335.152.678,58	

- Distribuição por Sistema de Concessão

Sistema de concessão	Descrição	Quantidade	Valor gerado (R\$)	% do Valor gerado
1	CONCESSAO FORMULARIO CCE	2	83.344,63	0,01%
4	CONCESSAO PRISMA/SUB	5.928	265.217.283,25	19,86%
5	CONCESSAO ON-LINE	45.888	764.414.791,20	57,25%
7	CONCESSAO DESCENTRALIZADA	2.534	109.310.733,03	8,19%
8	CONCESSAO ON-LINE, SEM INDICE	11.215	191.762.465,80	14,36%
11	CONCESSAO PELA INTERNET	217	4.364.060,67	0,33%
Total		65.784	1.335.152.678,58	

- Distribuição dos por forma de filiação (CS_FORMA_FILIAÇÃO)

Forma de Filiação	Descrição	Quantidade	Valor gerado (R\$)	% do Valor gerado
0	DESEMPREGADO	40.751	906.908.917,00	67,93%
1	EMPREGADO	10.850	263.492.592,35	19,74%
2	TRABALHADOR AVULSO	330	5.857.397,04	0,44%
3	EMPRESARIO	18	805.215,35	0,06%
4	DOMESTICO	1.684	17.693.165,71	1,33%
5	FACULTATIVO	1.620	18.190.004,83	1,36%
6	EQUIPARADO A AUTONOMO	9	267.596,63	0,02%
7	SEGURADO ESPECIAL	2.107	24.935.856,19	1,87%
8	AUTONOMO	8.346	96.335.117,59	7,22%
9	OPTANTE PELA LEI 6.184/74	69	666.815,89	0,05%
Total		65.784	1.335.152.678,58	

- Distribuição dos por forma de filiação (CS_RAMO_ATIVIDADE)

Ramo de Atividade	Descrição	Quantidade	Valor gerado (R\$)	% do Valor gerado
1	BANCARIO	74	6.721.800,25	0,50%
2	COMERCARIO	53.695	948.088.407,38	71,01%
3	TRANSPORTES E CARGA	824	19.996.450,10	1,50%
4	FERROVIARIO	19	803.413,40	0,06%
5	INDUSTRIARIO	7.279	296.119.703,54	22,18%
6	MARITIMO	15	738.225,24	0,06%
7	SERVIDOR PUBLICO	1.148	23.039.332,70	1,73%
8	RURAL	2.364	32.645.081,84	2,45%
9	IRRELEVANTE	366	7.000.264,13	0,52%
Total		65.784	1.335.152.678,58	

- Distribuição por Acesso ao CNIS (CS_BATIMENTO_CNIS)

Tipo de Acesso	Descrição	Quantidade	Valor gerado (R\$)	% do Valor gerado
*	NAO HOUVE UTILIZACAO DE DADOS DO CNIS	141	11.932.525,66	0,89%
000	NAO HOUVE UTILIZACAO DE DADOS DO CNIS	58.920	1.211.741.268,30	90,76%
001	ALTERACAO NOS DADOS BASICOS	934	14.060.540,87	1,05%
010	ALTERACAO NOS VINCULOS	216	3.553.908,18	0,27%
011	ALTER DADOS BASICOS, ALTERACAO VINCULOS	1.042	18.682.774,35	1,40%
020	INCLUSAO NOS VINCULOS	117	2.019.092,49	0,15%
021	ALTER DADOS BASICOS, INCLUSAO VINCULOS	450	7.649.499,38	0,57%
030	INCLUSAO/ALTERACAO NOS VINCULOS	134	2.176.311,37	0,16%
031	ALTER DADOS BASICOS, INCL/ALTER VINCULOS	523	9.348.125,82	0,70%
100	ALTERACAO NAS REMUNERACOES	3	28.218,61	0,00%
101	ALTER DADOS BASICOS, ALTER REMUNERACOES	4	81.144,31	0,01%
110	ALTERACAO VINCULOS, ALTER REMUNERACOES	1	21.417,16	0,00%
111	ALT DADOS BAS, ALT VINCULOS, ALT REMUNER	6	89.748,00	0,01%
120	INCLUSAO VINCULOS, ALTER REMUNERACOES	4	46.284,05	0,00%
121	ALT DADOS BAS, INC VINCULOS, ALT REMUNER	3	37.751,32	0,00%
131	ALT DADOS BASIC, INC/ALT VINC, ALT REMUN	5	60.179,42	0,00%
200	INCLUSAO NAS REMUNERACOES	9	145.771,46	0,01%
201	ALTER DADOS BASICOS, INCL REMUNERACOES	23	460.793,15	0,03%
210	ALTERACAO VINCULOS, INCLUSAO REMUNER	6	96.712,01	0,01%
211	ALT DADOS BAS, ALT VINCULOS, INC REMUNER	34	580.794,15	0,04%
220	INCLUSAO VINCULOS, INCLUSAO REMUNERACOES	7	86.904,68	0,01%
221	ALT DADOS BAS, INCL VINCULOS, INCL REMUN	22	342.452,37	0,03%
230	INCL/ALTER VINCULOS, INCL REMUNERACOES	6	119.187,59	0,01%
231	ALT DADOS BAS, INC/ALTER VINC, INC REMUN	17	274.865,63	0,02%
300	INCLUSAO/ALTERACAO NAS REMUNERACOES	3	41.257,47	0,00%
301	ALT DADOS BASICOS, INC/ALT REMUNERACOES	4	61.572,11	0,00%
310	ALT VINCULOS, INC/ALT NAS REMUNERACOES	7	99.360,26	0,01%
311	ALT DADOS BAS, ALT VINC, INC/ALT REMUNER	5	69.622,68	0,01%
320	INCLUSAO VINCULOS, INCL/ALT REMUNERACOES	3	28.841,47	0,00%
321	ALT DADOS BAS, INC VINC, INC/ALT REMUNER	9	155.597,31	0,01%
330	INC/ALT VINCULOS, INC/ALT REMUNERACOES	4	47.729,52	0,00%
331	ALT DADOS BAS, INC/ALT VINC, INC/ALT REM	3	42.455,05	0,00%
500	HOUVE UTIL. DADOS DO CNIS, SEM ALTERACAO	3.119	50.969.972,38	3,82%
Total		65.784	1.335.152.678,58	

- Distribuição pelas Regiões Geográficas Brasileiras

Unidade Geográfica	Quantidade	Valor gerado (R\$)	% do Valor gerado
NÃO INFORMADA	1	32.123,37	0,00%
CENTRO-OESTE	2.382	38.806.328,14	2,91%
NORDESTE	15.148	274.824.842,34	20,58%
NORTE	3.298	50.603.544,49	3,79%
SUDESTE	40.349	893.833.021,47	66,95%
SUL	4.606	77.052.818,77	5,77%
Total	65.784	1.335.152.678,58	

- Distribuição pela Quantidade de revisões (QT_REVISAO_RMI)

Quantidade de Revisões	Quantidade de Benefícios	Valor gerado (R\$)	% do Valor gerado
0	64.167	1.293.112.079,84	96,85%
1	1.531	37.948.029,30	2,84%
2	65	2.855.567,09	0,21%
3	9	445.723,62	0,03%
4	7	492.896,83	0,04%
5	1	26.588,00	0,00%
6	2	127.214,54	0,01%
7	1	70.783,44	0,01%
8	1	73.795,92	0,01%
Total	65.784	1.335.152.678,58	

9.6 APÊNDICE F – Análise dos resultados entre 99 e 100%

- Distribuição dos Benefícios por Espécie (CS_ESPECIE)

Espécie	Descrição	Quantidade	Valor gerado (R\$)	% do Valor gerado
31	AUXILIO DOENCA PREVIDENCIARIO	1.015	45.188.806,61	41,09%
32	APOSENTADORIA INVALIDEZ PREVIDENCIARIA	1.160	38.398.010,67	34,91%
91	AUXILIO DOENCA POR ACIDENTE DO TRABALHO	270	16.132.925,06	14,67%
92	APOSENT. INVALIDEZ ACIDENTE TRABALHO	152	10.267.195,93	9,33%
Total		2.597	109.986.938,27	

- Distribuição por Sistema de Concessão

Sistema de concessão	Descrição	Quantidade	Valor gerado (R\$)	% do Valor gerado
1	CONCESSAO FORMULARIO CCE	1	52.018,26	0,05%
4	CONCESSAO PRISMA/SUB	52	2.367.143,83	2,15%
5	CONCESSAO ON-LINE	44	74.092,00	0,07%
7	CONCESSAO DESCENTRALIZADA	2.492	107.441.276,60	97,69%
8	CONCESSAO ON-LINE, SEM INDICE	6	34.518,00	0,03%
11	CONCESSAO PELA INTERNET	2	17.889,58	0,02%
Total		2.597	109.986.938,27	

- Distribuição dos por forma de filiação (CS_FORMA_FILIAÇÃO)

Forma de Filiação	Descrição	Quantidade	Valor gerado (R\$)	% do Valor gerado
0	DESEMPREGADO	637	24.110.939,43	21,92%
1	EMPREGADO	1.220	64.272.785,03	58,44%
2	TRABALHADOR AVULSO	2	56.133,91	0,05%
3	EMPRESARIO	18	805.215,35	0,73%
4	DOMESTICO	52	1.528.198,07	1,39%
5	FACULTATIVO	39	1.099.589,09	1,00%
6	EQUIPARADO A AUTONOMO	5	206.810,26	0,19%
7	SEGURADO ESPECIAL	366	10.420.013,60	9,47%
8	AUTONOMO	257	7.458.553,53	6,78%
9	OPTANTE PELA LEI 6.184/74	1	28.700,00	0,03%
Total		2.597	109.986.938,27	

- Distribuição dos por forma de filiação (CS_RAMO_ATIVIDADE)

Ramo de Atividade	Descrição	Quantidade	Valor gerado (R\$)	% do Valor gerado
1	BANCARIO	34	3.895.337,64	3,54%
2	COMERCARIO	1.260	50.129.176,55	45,58%
3	TRANSPORTES E CARGA	67	4.338.463,24	3,94%
4	FERROVIARIO	3	266.860,92	0,24%
5	INDUSTRIARIO	549	29.129.428,77	26,48%
6	MARITIMO	1	94.605,44	0,09%
7	SERVIDOR PUBLICO	63	3.608.956,48	3,28%
8	RURAL	594	17.758.810,27	16,15%
9	IRRELEVANTE	26	765.298,96	0,70%
Total		2.597	109.986.938,27	

- Distribuição por Acesso ao CNIS (CS_batimento_CNIS)

Tipo de Acesso	Descrição	Quantidade	Valor gerado (R\$)	% do Valor gerado
000	NAO HOUVE UTILIZACAO DE DADOS DO CNIS	2.594	109.958.728,69	99,97%
020	INCLUSAO NOS VINCULOS	1	11.454,42	0,01%
500	HOUVE UTIL. DADOS DO CNIS, SEM ALTERACAO	2	16.755,16	0,02%
Total		2.597	109.986.938,27	

- Distribuição pelas Regiões Geográficas Brasileiras

Unidade Geográfica	Quantidade	Valor gerado (R\$)	% do Valor gerado
CENTRO-OESTE	50	2.521.113,24	2,29%
NORDESTE	428	15.897.637,75	14,45%
NORTE	92	3.442.118,99	3,13%
SUDESTE	1.551	68.427.051,10	62,21%
SUL	476	19.699.017,19	17,91%
Total	2.597	109.986.938,27	

- Distribuição pela Quantidade de revisões (QT_REVISAO_RMI)

Quantidade de Revisões	Quantidade	Valor gerado (R\$)	% do Valor gerado
0	2.312	96.357.573,23	0,876081967
1	219	9.875.177,47	0,089785002
2	47	2.639.701,89	0,02400014
3	8	397.002,87	0,003609546
4	7	492.896,83	0,004481412
5	1	26.588,00	0,000241738
6	2	127.214,54	0,001156633
7	1	70.783,44	0,000643562
Total	2.597	109.986.938,27	

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)