

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL

CESAR VIANNA MOREIRA JÚNIOR

**BUSCA PELO EQUILÍBRIO DA CARGA DE EXAME: UMA METODOLOGIA
PARA AVALIAÇÃO E DISTRIBUIÇÃO DE PEDIDOS DE PATENTE**

Rio de Janeiro

2021

M838 Moreira Júnior, Cesar Vianna. .

Busca pelo equilíbrio da carga de exame: uma metodologia para avaliação e distribuição de pedidos de patente. / Cesar Vianna Moreira Júnior. -- 2021.

187 f.; Fig. .

Tese (Doutorado em Propriedade Intelectual e Inovação) - Academia de Propriedade Intelectual Inovação e Desenvolvimento, Divisão de Programas de Pós-Graduação e Pesquisa, Instituto Nacional da Propriedade Industrial – INPI, Rio de Janeiro, 2021.

Orientador: Prof. Dr. Ricardo Carvalho Rodrigues.

Coorientador: Prof. Dr. Daniel Marques Golodne.

1. Propriedade industrial - Brasil. 2. Patente - Procedimento administrativo.
3. Patente - Exame. 4. Patente - Examinador. 5. Patente - Avaliação de carga.
- I. Instituto Nacional da Propriedade Industrial (Brasil).

CDU: 347.771:658(81)

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta dissertação, desde que citada a fonte.

Assinatura

Data

Cesar Vianna Moreira Júnior

Busca pelo equilíbrio da carga de exame: uma metodologia para avaliação e distribuição de pedidos de patente

Tese apresentada, como requisito parcial para obtenção do título de Doutor, ao Programa de Pós-Graduação em Propriedade Intelectual e Inovação, do Instituto Nacional da Propriedade Industrial.

Orientador: Dr. Ricardo Carvalho Rodrigues

Coorientador: Dr. Daniel Marques Golodne

Rio de Janeiro

2021

Cesar Vianna Moreira Júnior

**Busca pelo equilíbrio da carga de exame: uma metodologia para avaliação e distribuição
de pedidos de patente**

Tese apresentada, como requisito parcial para obtenção do título de Doutor, ao Programa de Pós-Graduação em Propriedade Intelectual e Inovação, do Instituto Nacional da Propriedade Industrial.

Aprovada em 25 de Junho de 2021

Orientador: Prof. Dr. Ricardo Carvalho Rodrigues
Instituto Nacional da Propriedade Industrial

Coorientador: Dr. Daniel Marques Golodne
Instituto Nacional da Propriedade Industrial

Banca Examinadora:
Dr. Alexandre Gomes Ciancio
Instituto Nacional da Propriedade Industrial
Dr. Daniel de Souza Dias
Instituto Nacional da Propriedade Industrial
Prof^a. Dra^a. Adelaide Maria de Souza Antunes
Instituto Nacional da Propriedade Industrial
Prof. Dr. Alexandre Guimarães Vasconcellos
Instituto Nacional da Propriedade Industrial

A ata de defesa com respectivas assinaturas dos membros da banca examinadora encontra-se no processo de vida acadêmica do aluno.

Rio de Janeiro

2021

DEDICATÓRIA

Dedico este trabalho à minha esposa, Fabíola Holzbecher Fragozo, pelo apoio, compreensão e incentivo durante todo este longo processo, e aos meus pais Cesar Vianna Moreira e Maria Cristina Matarazzo Freire Moreira, pela dedicação e suporte em todas as etapas da minha formação pessoal e profissional.

AGRADECIMENTOS

Agradeço a todas as pessoas que, direta e/ou indiretamente, contribuíram para a execução e finalização desta etapa em minha vida acadêmica. No entanto, alguns merecem um agradecimento especial.

Inicialmente, agradeço ao meu orientador Dr. Ricardo Carvalho Rodrigues que, acreditando na proposta de pesquisa e em minha capacidade de desenvolvê-la, e com bastante tolerância, observações críticas e construtivas, contribuiu decisivamente para o desenvolvimento das soluções e resultados alcançados. Tenho certeza de que ainda vamos fazer muitos trabalhos juntos e que, além de orientador, tornou-se um grande parceiro.

Agradeço ao Dr. Daniel Marques Golodne, meu coorientador, pela disponibilidade, interesse e colaboração em solucionar problemas e dúvidas, além da contribuição com sugestão de relevantes artigos técnicos utilizados no trabalho. Sua participação também foi importante para o bom andamento desta tese.

Aos professores e funcionários do Programa de Pós Graduação do Instituto Nacional da Propriedade Industrial, meu reconhecimento e profundo agradecimento.

Aos professores Dr.^a Adelaide Maria de Souza Antunes, Dr. Alexandre Guimarães Vasconcellos e Dr. Celso Salgueiro Lage, membros da banca de qualificação, pelas sugestões.

Aos membros da banca examinadora, Dr. Alexandre Gomes Ciancio, Dr. Daniel de Souza Dias e professores Dr.^a Adelaide Maria de Souza Antunes e Dr. Alexandre Guimarães Vasconcellos, que, gentilmente, aceitaram participar e colaborar com esta tese.

Aos novos amigos que a academia me proporcionou, em particular, Deuzanira Santos, Jaqueline Carolino e Maria Ângela Fernandes que, além de toda sinergia, cooperação e diversas parcerias de produção intelectual, deram apoio e atenção fundamentais nos momentos mais difíceis da minha caminhada.

Não há palavras que possam externar toda gratidão aos amigos Pedro Leal de Lima Soares e Luiz Claudio Teixeira Herig que, com sua disponibilidade, paciência, gentileza e conhecimento técnico, foram essenciais para o desenvolvimento e conclusão deste trabalho. Destaco que este agradecimento se estende ao meu chefe Bernardo Seelig e a todos os meus colegas de divisão, pelos debates, incentivo ou participação em etapas do meu trabalho, em especial aos examinadores Alexandre Bissoli, Bruno Montesano e Valmir Junior.

E, finalmente, minha família que, seguramente, compartilha da minha alegria.

EPÍGRAFE

A menos que modifiquemos a nossa maneira de pensar, não seremos capazes de resolver os problemas causados pela forma como nos acostumamos a ver o mundo.

Albert Einstein

MOREIRA JÚNIOR, Cesar Vianna. **Busca pelo equilíbrio da carga de exame: uma metodologia para avaliação e distribuição de pedidos de patente.** Rio de Janeiro, 2021. 187. f. Tese (Doutorado em Propriedade Intelectual e Inovação) – Instituto Nacional da Propriedade Industrial, Rio de Janeiro, 2021.

RESUMO

Esta tese tem como principal objetivo apresentar uma metodologia para avaliação e distribuição de pedidos de patente aos examinadores de patente do Instituto Nacional da Propriedade Industrial - INPI levando em consideração não somente o campo tecnológico específico, representado pelo enquadramento do pedido na Classificação Internacional de Patentes (CIP), mas também variáveis relacionadas ao volume de dados do pedido e sua complexidade para o processo de exame. O desenvolvimento proposto pode ser dividido em seis etapas, que se complementam entre si: (i) identificação das variáveis de interesse por meio da avaliação da estrutura básica de um pedido de patente com a análise de documentos reais, revisão do arcabouço legal específico (leis, diretrizes, procedimentos, instruções normativas, guias e artigos) e uma revisão bibliográfica de estudos que utilizam como base variáveis de documentos de patente; (ii) pesquisa de métodos de análise multivariada para utilização como ferramenta auxiliar de avaliação dos pedidos de patente quanto ao volume de dados e complexidade, e seleção do método de análise de componentes principais (ACP); (iii) desenvolvimento de um modelo matemático compreendendo o cálculo de um Índice Geral de Complexidade (IGC) para cada um dos pedidos de patente, enquadramento dos pedidos em cinco classes (muito leve, leve, moderado, pesado e muito pesado) conforme faixas do IGC e uma lógica de distribuição compensando pedidos muito pesados com muito leves e leves com pesados; (iv) avaliação da lógica de distribuição por meio do cálculo de um Índice de Balanceamento da Distribuição (IBD) considerando as diferenças das medianas das amostras e da realização de testes de hipótese de Mann-Whitney; (v) realização de testes de validação do modelo por meio de uma amostra de pedidos da área de eletricidade incluindo, além das variáveis identificadas, o tempo dispendido no exame substantivo por um examinador de referência; e (vi) realização de testes de aplicabilidade com amostras de outras áreas. Os resultados sugerem que as variáveis relevantes a serem consideradas são o número total de páginas, número de reivindicações e número de páginas de reivindicações do pedido, sendo que esta última apresentou o comportamento mais estável de correlações com o tempo, além de valores elevados (cerca de 0,75). Adicionalmente, os resultados dos testes de validação mostraram altas correlações do IGC com o tempo, acima de 80%, acerto nos enquadramentos também acima de 80%, aumentos de até 8,5% do IBD, e aprovação das amostras nos testes de Mann-Whitney, ou seja, o modelo matemático conseguiu representar de maneira bastante satisfatória o esforço/tempo de exame de pedidos de patente com aproveitamento de buscas de escritórios internacionais. Vale ressaltar que, das simulações realizadas, considera-se que o modelo pode ser aplicado a outras áreas do INPI desde que seja realizada uma análise estatística prévia das amostras e sejam atendidos critérios mínimos relativos ao tamanho da amostra, variância acumulada das componentes principais, distribuição estatística dos dados do IGC, correlação dos dados de IGC com o número de páginas de reivindicações e valores IBD.

Palavras-chave: pedidos de patente; avaliação; distribuição; volume de dados; CIP; tempo de exame; examinadores de patente; componentes principais.

MOREIRA JÚNIOR, Cesar Vianna. **Seeking to balance examiner's workload: a methodology for evaluation and distribution of patent applications**. Rio de Janeiro, 2021. 187. f. Tese (Doutorado em Propriedade Intelectual e Inovação) – Instituto Nacional da Propriedade Industrial, Rio de Janeiro, 2021.

ABSTRACT

This thesis has as main objective to present a methodology for the evaluation and distribution of patent applications to the examiners of the Brazilian Patent Office considering not only the specific technological field, represented by the application's framework in the International Patent Classification (IPC), but also variables related to the application's volume data and its complexity for the examination process. The proposed development was divided into six parts, which are complemented by one another: (i) identification of the relevant variables through the evaluation of the basic structure of a patent application with the analysis of real documents, review of the specific legal framework (laws, guidelines, procedures, normative instructions, guides and articles) and a bibliographic review of studies based on variables of patent documents; (ii) research on multivariate analysis methods for use as an auxiliary tool for assessing patent applications in terms of data volume and complexity, and selection of the principal component analysis method (PCA); (iii) development of a mathematical model including the calculation of a General Complexity Index (GCI) for each of the patent applications, classification of the applications in five classes (very light, light, moderate, heavy and very heavy) according to ranges of the GCI and a distribution logic compensating very heavy applications with very light and light with heavy ones; (iv) evaluation of the distribution logic by calculating a Distribution Balancing Index (DBI) considering the differences in the medians of the samples and the performance of Mann-Whitney hypothesis tests; (v) conducting model validation tests through a sample from the electricity area including, in addition to the variables identified, the substantive examination time by a reference examiner; and (vi) conducting tests with samples from other areas. The results suggest that the most relevant variables are the total number of pages, number of claims and number of pages of claims, the later one having the most stable behavior of correlations with time, and also high values (about 0.75). Additionally, the results of the validation tests showed IGC has high correlations with time, above 80%, a success in the model frameworks also above 80%, increases up to 8,5% of the IBD values and approval of the samples in the Mann-Whitney tests. Therefore, the mathematical model was able to represent satisfactorily the effort/time of patent examination using the search reports from international offices. It is worth mentioning that, from the simulations carried out, it is considered that the model can be applied to other areas of the INPI provided that a prior statistical analysis of the samples is carried out and minimum criteria are met regarding the sample size, accumulated variance of the principal components, IGC statistical distribution, IGC correlation with the number of claim pages and IBD values.

Keywords: patent applications; evaluation; distribution; volume data; IPC; examination time; patent examiners; principal components.

SUMÁRIO

1 INTRODUÇÃO	16
1.1 OBJETIVOS.....	17
1.2 ORGANIZAÇÃO DO TEXTO	19
2 O DOCUMENTO DE PATENTE E AS VARIÁVEIS DE INTERESSE PARA AVALIAÇÃO E DISTRIBUIÇÃO	21
2.1 ESTRUTURA BÁSICA DO PEDIDO DE PATENTE.....	21
2.1.1 RELATÓRIO DESCRITIVO	21
2.1.2 REIVINDICAÇÕES.....	22
2.1.3 DESENHOS.....	23
2.1.4 RESUMO	24
2.2 FOLHA DE ROSTO: DADOS BIBLIOGRÁFICOS DO DOCUMENTO DE PATENTE	25
2.3 CLASSIFICAÇÃO INTERNACIONAL DE PATENTES (CIP).....	26
2.4 SUBSÍDIOS AO EXAME TÉCNICO.....	28
2.5 ESTUDOS UTILIZANDO VARIÁVEIS DOS DOCUMENTOS DE PATENTE	29
2.5.1 ESTUDOS DIVERSOS.....	29
2.5.2 BALANCEAMENTO DE CARGA, VOLUMINOSIDADE E COMPLEXIDADE DOS PEDIDOS DE PATENTE.....	30
2.6 VARIÁVEIS SELECIONADAS E CONSIDERAÇÕES INICIAIS.....	31
3 METODOLOGIA	33
3.1 IDENTIFICAÇÃO DAS VARIÁVEIS DE INTERESSE	33
3.2 AVALIAÇÃO DOS PEDIDOS DE PATENTE QUANTO AO VOLUME DE DADOS E COMPLEXIDADE	34
3.3 DISTRIBUIÇÃO DOS PEDIDOS DE PATENTE	36
3.4 AVALIAÇÃO DA NOVA LÓGICA DE DISTRIBUIÇÃO	37
3.5 AVALIAÇÃO DA APLICABILIDADE EM OUTRAS ÁREAS DO INPI	38
4 DESENVOLVIMENTO	39
4.1 ANÁLISE MULTIVARIADA: APLICAÇÃO DA ANÁLISE DAS COMPONENTES PRINCIPAIS (ACP).....	39
4.1.1 INTRODUÇÃO.....	39
4.1.2 DETERMINAÇÃO DAS COMPONENTES PRINCIPAIS.....	40
4.1.2.1 PADRONIZAÇÃO DAS VARIÁVEIS	41
4.1.2.2 MATRIZES DE VARIÂNCIA-COVARIÂNCIA (S) E DE CORRELAÇÃO (R)	42
4.1.2.3 CÁLCULO DOS AUTOVALORES E AUTOVETORES DAS MATRIZES S E R	43
4.1.2.4 CÁLCULO DAS COMPONENTES PRINCIPAIS E ALGUMAS PROPRIEDADES	44
4.1.2.5 CRITÉRIOS DE SELEÇÃO DAS COMPONENTES PRINCIPAIS.....	46
4.1.3 ADEQUAÇÃO DOS DADOS.....	47
4.1.3.1 NÚMERO DE VARIÁVEIS E ELEMENTOS DAS AMOSTRAS	47
4.1.3.2 TESTE DE BARTLETT	47
4.1.3.3 TESTE DE KAISER-MEYER-OLKIN (KMO)	48
4.1.4 ÍNDICE GERAL DE COMPLEXIDADE (IGC)	49
4.1.5 ENQUADRAMENTO DE CLASSES DOS PEDIDOS DE PATENTE	49
4.2 EXAME SUBSTANTIVO DE PATENTES	52
4.2.1 EXAME TÉCNICO – ETAPAS PRINCIPAIS, PARECER TÉCNICO E VARIÁVEIS RELACIONADAS.....	52
4.2.1.1 ETAPAS PRINCIPAIS.....	52
4.2.1.1.1 ETAPA 1 – ANÁLISE TÉCNICA INICIAL	52
4.2.1.1.2 ETAPA 2 – BUSCA DE ANTERIORIDADES	53
4.2.1.1.3 ETAPA 3 – ANÁLISE TÉCNICA ESPECÍFICA	53
4.2.1.2 PARECER TÉCNICO DO INPI – ESTRUTURA E PREENCHIMENTO	53
4.2.2 FORMULÁRIO DE TEMPO DE EXAME.....	57
4.3 NOVA METODOLOGIA PARA DISTRIBUIÇÃO DOS PEDIDOS DE PATENTE	58
4.3.1 DETERMINAÇÃO DAS ZONAS DE ATUAÇÃO ESPECÍFICAS (ZAE) DOS EXAMINADORES	58
4.3.2 LÓGICA ITERATIVA PARA DISTRIBUIÇÃO INDIVIDUAL DOS PEDIDOS.....	58

4.3.3 AMOSTRA FINAL DE REDISTRIBUIÇÃO E CÁLCULO DO ÍNDICE DE BALANCEAMENTO DA DISTRIBUIÇÃO (IBD) 60

5 RESULTADOS	62
5.1 AMOSTRA DE TESTE INICIAL	63
5.1.1 VARIÁVEIS DE INTERESSE ANALISADAS	63
5.1.2 ANÁLISE DA AMOSTRA DE TESTE INICIAL	64
5.1.3 EXECUÇÃO DA ANÁLISE DE COMPONENTES PRINCIPAIS	67
5.1.3.1 PADRONIZAÇÃO DA AMOSTRA DE TESTE INICIAL	67
5.1.3.2 CÁLCULO DA MATRIZ DE CORRELAÇÃO E SELEÇÃO DAS COMPONENTES PRINCIPAIS	68
5.1.3.3 TESTES DE ADEQUAÇÃO DOS DADOS	69
5.1.3.3.1 TESTE DE BARTLETT	69
5.1.3.3.2 TESTE KMO	70
5.1.3.4 CÁLCULO E AVALIAÇÃO DAS COMPONENTES PRINCIPAIS	70
5.1.3.5 AVALIAÇÃO DOS PEDIDOS DE PATENTE DA AMOSTRA DE TESTE INICIAL	72
5.1.4 ANÁLISE DE SUBCLASSES DA CIP	77
5.1.4.1 ZAE DOS EXAMINADORES	77
5.1.4.2 AVALIAÇÃO DAS SUBCLASSES DA CIP DAS ZAE	78
5.1.4.3 REDISTRIBUIÇÃO DOS PEDIDOS	80
5.1.4.3.1 NOVA CONFIGURAÇÃO DAS ZAE DOS EXAMINADORES	80
5.1.4.3.2 NOVA CONFIGURAÇÃO DOS PEDIDOS REDISTRIBUÍDOS	81
5.1.4.3.3 ÍNDICE DE BALANCEAMENTO DA DIVISÃO (IBD) ANTES E APÓS A REDISTRIBUIÇÃO	82
5.2 AMOSTRA PADRÃO COM TEMPO	82
5.2.1 COMPOSIÇÃO DA AMOSTRA E VARIÁVEIS DE INTERESSE	82
5.2.2 SIMULAÇÕES E TESTES: ANÁLISE DE SENSIBILIDADE DO MODELO	85
5.2.2.1 CASOS DE ESTUDO: APLICAÇÃO DA ACP E ANÁLISE DE CORRELAÇÕES	85
5.2.2.1.1 CASO 1: CASO 9 VAR	85
5.2.2.1.1.1 CÁLCULO DA MATRIZ DE CORRELAÇÃO E SELEÇÃO DAS COMPONENTES PRINCIPAIS	86
5.2.2.1.1.2 CÁLCULO E AVALIAÇÃO DAS COMPONENTES PRINCIPAIS	88
5.2.2.1.1.3 CORRELAÇÕES COM TEMPO	89
5.2.2.1.2 CASO 2: CASO 5 VAR	90
5.2.2.1.2.1 CÁLCULO DA MATRIZ DE CORRELAÇÃO E SELEÇÃO DAS COMPONENTES PRINCIPAIS	91
5.2.2.1.2.2 CÁLCULO E AVALIAÇÃO DAS COMPONENTES PRINCIPAIS	93
5.2.2.1.2.3 CORRELAÇÕES COM O TEMPO	93
5.2.2.1.3 CASO 3: CASO 4 VAR	95
5.2.2.1.3.1 CÁLCULO DA MATRIZ DE CORRELAÇÃO E SELEÇÃO DAS COMPONENTES PRINCIPAIS	95
5.2.2.1.3.2 CÁLCULO E AVALIAÇÃO DAS COMPONENTES PRINCIPAIS	98
5.2.2.1.3.3 CORRELAÇÕES COM TEMPO	98
5.2.2.1.4 CASO 4: CASO 3 VAR	99
5.2.2.1.4.1 CÁLCULO DA MATRIZ DE CORRELAÇÃO E SELEÇÃO DAS COMPONENTES PRINCIPAIS	100
5.2.2.1.4.2 CÁLCULO E AVALIAÇÃO DAS COMPONENTES PRINCIPAIS	101
5.2.2.1.4.3 CORRELAÇÕES COM O TEMPO	101
5.2.2.1.5 CASO 5: CASO 3 VAR (2)	102
5.2.2.1.5.1 CÁLCULO DA MATRIZ DE CORRELAÇÃO E SELEÇÃO DAS COMPONENTES PRINCIPAIS	103
5.2.2.1.5.2 CÁLCULO E AVALIAÇÃO DAS COMPONENTES PRINCIPAIS	105
5.2.2.1.5.3 CORRELAÇÕES COM TEMPO	105
5.2.2.2 ANÁLISE DE SENSIBILIDADE DO IGC: COMPARAÇÃO DOS CASOS	106
5.2.2.3 ENQUADRAMENTO EM CLASSES DOS PEDIDOS DE PATENTE DA AMOSTRA PADRÃO	112
5.2.2.3.1 ENQUADRAMENTO DOS PEDIDOS DE PATENTE DA AMOSTRA PADRÃO POR TEMPO	112
5.2.2.3.2 ENQUADRAMENTO DOS PEDIDOS DE PATENTE DA AMOSTRA PADRÃO POR IGC _{Y1}	115
5.3 TESTE DE REDISTRIBUIÇÃO – ÁREA DE ELETRICIDADE	121
5.3.1 AMOSTRA FINAL DE REDISTRIBUIÇÃO	121
5.3.2 EXECUÇÃO DA ANÁLISE DE COMPONENTES PRINCIPAIS: CASO 3 VAR (2)	122
5.3.3 ENQUADRAMENTO DOS PEDIDOS DE PATENTE DA AMOSTRA FINAL DE REDISTRIBUIÇÃO POR IGC _{Y1}	123
5.3.4 ANÁLISE DAS SUBCLASSES DA CIP DA AMOSTRA FINAL DE REDISTRIBUIÇÃO	127
5.3.4.1 ZAE ORIGINAL DOS EXAMINADORES DA AMOSTRA FINAL DE REDISTRIBUIÇÃO	127
5.3.4.2 CONFIGURAÇÃO ORIGINAL DOS PEDIDOS	128
5.3.4.3 REDISTRIBUIÇÃO DOS PEDIDOS	129
5.3.4.3.1 NOVA CONFIGURAÇÃO DAS ZAE DOS EXAMINADORES	129
5.3.4.3.2 NOVA CONFIGURAÇÃO DOS PEDIDOS REDISTRIBUÍDOS	130

5.3.4.3.3	ÍNDICE DE BALANCEAMENTO DA DISTRIBUIÇÃO (IBD) ANTES E APÓS A REDISTRIBUIÇÃO	131
5.3.4.3.4	TESTES DE HIPÓTESE	131
5.4	TESTES DE REDISTRIBUIÇÃO – OUTRAS ÁREAS	132
5.4.1	AMOSTRA DE REDISTRIBUIÇÃO – ÁREA DE ELETRÔNICA	133
5.4.1.1	EXECUÇÃO DA ANÁLISE DE COMPONENTES PRINCIPAIS: CASO 3 VAR (2)	133
5.4.1.2	ENQUADRAMENTO DOS PEDIDOS DE PATENTE DA AMOSTRA DE REDISTRIBUIÇÃO POR IGC_{Y1}	135
5.4.1.3	ANÁLISE DAS SUBCLASSES DA CIP DA AMOSTRA DE REDISTRIBUIÇÃO	138
5.4.1.3.1	ZAE ORIGINAL DOS EXAMINADORES DA AMOSTRA DE REDISTRIBUIÇÃO	138
5.4.1.3.2	CONFIGURAÇÃO ORIGINAL DOS PEDIDOS	139
5.4.1.3.3	REDISTRIBUIÇÃO DOS PEDIDOS	140
5.4.1.3.3.1	NOVA CONFIGURAÇÃO DAS ZAE DOS EXAMINADORES	140
5.4.1.3.3.2	NOVA CONFIGURAÇÃO DOS PEDIDOS REDISTRIBUÍDOS	141
5.4.1.3.3.3	ÍNDICE DE BALANCEAMENTO DA DISTRIBUIÇÃO (IBD) ANTES E APÓS A REDISTRIBUIÇÃO	142
5.4.1.3.3.4	TESTES DE HIPÓTESE	142
5.4.2	AMOSTRA DE REDISTRIBUIÇÃO – ÁREA DE QUÍMICA (POLÍMEROS)	143
5.4.2.1	EXECUÇÃO DA ANÁLISE DE COMPONENTES PRINCIPAIS: CASO 3 VAR (2)	143
5.4.2.2	ENQUADRAMENTO DOS PEDIDOS DE PATENTE DA AMOSTRA FINAL DE REDISTRIBUIÇÃO POR IGC_{Y1} 145	
5.4.2.3	ANÁLISE DAS SUBCLASSES DA CIP DA AMOSTRA DE REDISTRIBUIÇÃO	148
5.4.2.3.1	ZAE ORIGINAL DOS EXAMINADORES DA AMOSTRA DE REDISTRIBUIÇÃO	148
5.4.2.3.2	CONFIGURAÇÃO ORIGINAL DOS PEDIDOS	148
5.4.2.3.3	REDISTRIBUIÇÃO DOS PEDIDOS	149
5.4.2.3.3.1	NOVA CONFIGURAÇÃO DAS ZAE DOS EXAMINADORES	149
5.4.2.3.3.2	NOVA CONFIGURAÇÃO DOS PEDIDOS REDISTRIBUÍDOS	150
5.4.2.3.3.3	ÍNDICE DE BALANCEAMENTO DA DISTRIBUIÇÃO (IBD) ANTES E APÓS A REDISTRIBUIÇÃO	151
5.4.2.3.3.4	TESTES DE HIPÓTESE	151
5.4.3	AMOSTRA DE REDISTRIBUIÇÃO – ÁREA DE BIOLOGIA (BIOFÁRMACOS)	152
5.4.3.1	EXECUÇÃO DA ANÁLISE DE COMPONENTES PRINCIPAIS: CASO 3 VAR (2)	153
5.4.3.2	ENQUADRAMENTO DOS PEDIDOS DE PATENTE DA AMOSTRA DE REDISTRIBUIÇÃO POR IGC_{Y1}	155
5.4.3.3	ANÁLISE DAS SUBCLASSES DA CIP DA AMOSTRA DE REDISTRIBUIÇÃO	157
5.4.3.3.1	ZAE ORIGINAL DOS EXAMINADORES DA AMOSTRA DE REDISTRIBUIÇÃO	157
5.4.3.3.2	CONFIGURAÇÃO ORIGINAL DOS PEDIDOS	158
5.4.3.3.3	REDISTRIBUIÇÃO DOS PEDIDOS	159
5.4.3.3.3.1	NOVA CONFIGURAÇÃO DAS ZAE DOS EXAMINADORES	159
5.4.3.3.3.2	NOVA CONFIGURAÇÃO DOS PEDIDOS REDISTRIBUÍDOS	159
5.4.3.3.3.3	ÍNDICE DE BALANCEAMENTO DA DISTRIBUIÇÃO (IBD) ANTES E APÓS A REDISTRIBUIÇÃO	160
5.4.3.3.3.4	TESTES DE HIPÓTESE	161
5.4.4	AMOSTRA DE REDISTRIBUIÇÃO – ÁREA DE MECÂNICA	161
5.4.4.1	EXECUÇÃO DA ANÁLISE DE COMPONENTES PRINCIPAIS: CASO 3 VAR (2)	162
5.4.4.2	ENQUADRAMENTO DOS PEDIDOS DE PATENTE DA AMOSTRA DE REDISTRIBUIÇÃO POR IGC_{Y1}	164
5.4.4.3	ANÁLISE DAS SUBCLASSES DA CIP DA AMOSTRA DE REDISTRIBUIÇÃO	166
5.4.4.3.1	ZAE ORIGINAL DOS EXAMINADORES DA AMOSTRA DE REDISTRIBUIÇÃO	166
5.4.4.3.2	CONFIGURAÇÃO ORIGINAL DOS PEDIDOS	167
5.4.4.3.3	REDISTRIBUIÇÃO DOS PEDIDOS	168
5.4.4.3.3.1	NOVA CONFIGURAÇÃO DAS ZAE DOS EXAMINADORES	168
5.4.4.3.3.2	NOVA CONFIGURAÇÃO DOS PEDIDOS REDISTRIBUÍDOS	169
5.4.4.3.3.3	ÍNDICE DE BALANCEAMENTO DA DISTRIBUIÇÃO (IBD) ANTES E APÓS A REDISTRIBUIÇÃO	170
5.4.4.3.3.4	TESTES DE HIPÓTESE	170
5.4.5	COMPARAÇÃO DE ÁREAS TÉCNICAS - AMOSTRAS DE REDISTRIBUIÇÃO	171
6	CONCLUSÃO	176
	REFERÊNCIAS	182

LISTA DE TABELAS

Tabela 1 – Matriz X de dados de n indivíduos (pedidos de patente) com p variáveis	40
Tabela 2 – Enquadramento de Classes dos Pedidos de Patente	50
Tabela 3 – Estrutura da Amostra de Teste Inicial	63
Tabela 4 – Estatísticas básicas da Amostra de Teste Inicial	64
Tabela 5 – Medianas da Amostra de Teste Inicial separadas por ano de depósito do pedido	65
Tabela 6 – Medianas da Amostra de Teste Inicial separadas por examinador	66
Tabela 7 – Estrutura da Amostra de Teste Inicial com dados padronizados	66
Tabela 8 – Autovalores e Contribuições de Variância das Componentes Principais	68
Tabela 9 – Escores das Componentes Principais de cada variável.....	69
Tabela 10 – Estrutura de dados das Componentes Principais selecionadas	70
Tabela 11 – Estrutura de dados das Componentes Principais selecionadas	71
Tabela 12 – Comparação de enquadramentos pelo IGC, total de páginas e total de reivindicações	74
Tabela 13 – Comparação de enquadramentos pelo IGC e pela componente Y1	75
Tabela 14 – Pedidos enquadrados como muito pesados por IGC	75
Tabela 15 – Zonas de Atuação dos Examinadores por Subclasses da CIP	76
Tabela 16 – Médias, Desvios Padrão e Coeficientes de Variação das variáveis por Subclasse e por Pedido	77
Tabela 17 – Coeficientes de Variação das Variáveis de cada Subclasse.....	78
Tabela 18 – Enquadramentos dos pedidos por IGC por subclasse.....	78
Tabela 19 – ZAE dos examinadores após a nova distribuição	79
Tabela 20 – Porcentagem de pedidos dentro das ZAE antes e após a nova distribuição	79
Tabela 21 – Medianas da Amostra de Teste Inicial redistribuída separadas por examinador	80
Tabela 22 – Formulário de tempo de exame (parte 1) preenchido	82
Tabela 23 – Formulário de tempo de exame (parte 2) preenchido	82
Tabela 24 – Estrutura da Amostra Padrão.....	83
Tabela 25 – Estatísticas básicas da Amostra Padrão	83
Tabela 26 – Autovalores e Contribuições de Variância das Componentes Principais	85
Tabela 27 – Escores das Componentes Principais de cada variável.....	87
Tabela 28 – Autovalores e Contribuições de Variância das Componentes Principais	90
Tabela 29 – Escores das Componentes Principais de cada variável.....	92
Tabela 30 – Autovalores e Contribuições de Variância das Componentes Principais	94
Tabela 31 – Escores das Componentes Principais de cada variável.....	97
Tabela 32 – Autovalores e Contribuições de Variância das Componentes Principais	99
Tabela 33 – Escores das Componentes Principais de cada variável.....	101
Tabela 34 – Autovalores e Contribuições de Variância das Componentes Principais	103
Tabela 35 – Escores das Componentes Principais de cada variável.....	105
Tabela 36 – Pedidos de patente enquadrados como leves por tempo.....	114
Tabela 37 – Pedidos de patente enquadrados como pesados por tempo	114
Tabela 38 – Pedido de patente enquadrado como muito pesado por tempo.....	115
Tabela 39 – Pedidos de patente enquadrados como leves por IGC_{Y1}	117
Tabela 40 – Pedidos de patente enquadrados como pesados por IGC_{Y1}	117
Tabela 41 – Pedido de patente enquadrado como muito pesado por IGC_{Y1}	117
Tabela 42 – Pedidos de patente com diferença nos enquadramentos por tempo e IGC_{Y1}	118
Tabela 43 – Estatísticas Básicas da Amostra Final de Redistribuição	122
Tabela 44 – Autovalores e Contribuições de Variância das Componentes Principais	123
Tabela 45 – Escores das Componentes Principais de cada variável.....	123
Tabela 46 – Pedidos de patente enquadrados como muito pesados por IGC_{Y1}	125
Tabela 47 – Pedidos de patente enquadrados como pesados por IGC_{Y1}	125
Tabela 48 – Pedidos de patente enquadrados como leves por IGC_{Y1}	125
Tabela 49 – Zonas de Atuação dos Examinadores por Subclasses da CIP	127
Tabela 50 – Medianas da Amostra de Final de Redistribuição separadas por examinador.....	128
Tabela 51 – ZAE dos examinadores após a nova distribuição	129
Tabela 52 – Porcentagem de pedidos dentro das ZAE antes e após a nova distribuição	129
Tabela 53 – Medianas da Amostra de Final de Redistribuição separadas por examinador.....	130
Tabela 54 – Resultados de p-valor dos Testes de Mann-Whitney	132
Tabela 55 – Estatísticas Básicas da Amostra de Redistribuição	133
Tabela 56 – Autovalores e Contribuições de Variância das Componentes Principais	133
Tabela 57 – Escores das Componentes Principais de cada variável.....	135

Tabela 58 – Pedidos de patente enquadrados como muito pesados por IGC_{Y1}	137
Tabela 59 – Pedidos de patente enquadrados como pesados por IGC_{Y1}	137
Tabela 60 – Pedidos de patente enquadrados como leves por IGC_{Y1}	137
Tabela 61 – Zonas de Atuação dos Examinadores por Subclasses da CIP	139
Tabela 62 – Medianas da Amostra de Redistribuição separadas por examinador.....	139
Tabela 63 – ZAE dos examinadores após a nova distribuição	140
Tabela 64 – Porcentagem de pedidos dentro das ZAE antes e após a nova distribuição	140
Tabela 65 – Medianas da Amostra de Redistribuição separadas por examinador.....	141
Tabela 66 – Resultados de p-valor dos Testes de Mann-Whitney	142
Tabela 67 – Estatísticas Básicas da Amostra de Redistribuição	143
Tabela 68 – Autovalores e Contribuições de Variância das Componentes Principais	143
Tabela 69 – Escores das Componentes Principais de cada variável.....	145
Tabela 70 – Pedido de patente enquadrado como muito pesado por IGC_{Y1}	146
Tabela 71 – Pedidos de patente enquadrados como pesados por IGC_{Y1}	146
Tabela 72 – Pedidos de patente enquadrados como leves por IGC_{Y1}	147
Tabela 73 – Zonas de Atuação dos Examinadores por Subclasses da CIP	148
Tabela 74 – Medianas da Amostra de Redistribuição separadas por examinador.....	149
Tabela 75 – ZAE dos examinadores após a nova distribuição	149
Tabela 76 – Porcentagem de pedidos dentro das ZAE antes e após a nova distribuição	150
Tabela 77 – Medianas da Amostra de Redistribuição separadas por examinador.....	150
Tabela 78 – Resultados de p-valor dos Testes de Mann-Whitney	151
Tabela 79 – Estatísticas Básicas da Amostra de Redistribuição	152
Tabela 80 – Autovalores e Contribuições de Variância das Componentes Principais	153
Tabela 81 – Escores das Componentes Principais de cada variável.....	154
Tabela 82 – Pedido de patente enquadrado como muito pesado por IGC_{Y1}	156
Tabela 83 – Pedidos de patente enquadrados como pesados por IGC_{Y1}	156
Tabela 84 – Pedidos de patente enquadrados como leves por IGC_{Y1}	156
Tabela 85 – Zonas de Atuação dos Examinadores por Subclasses da CIP	157
Tabela 86 – Medianas da Amostra de Redistribuição separadas por examinador.....	158
Tabela 87 – ZAE dos examinadores após a nova distribuição	159
Tabela 88 – Porcentagem de pedidos dentro das ZAE antes e após a nova distribuição	159
Tabela 89 – Medianas da Amostra de Redistribuição separadas por examinador.....	160
Tabela 90 – Resultados de p-valor dos Testes de Mann-Whitney	161
Tabela 91 – Estatísticas Básicas da Amostra de Redistribuição	161
Tabela 92 – Autovalores e Contribuições de Variância das Componentes Principais	162
Tabela 93 – Escores das Componentes Principais de cada variável.....	163
Tabela 94 – Pedido de patente enquadrado como muito pesado por IGC_{Y1}	165
Tabela 95 – Pedidos de patente enquadrados como pesados por IGC_{Y1}	165
Tabela 96 – Pedidos de patente enquadrados como leves por IGC_{Y1}	165
Tabela 97 – Zonas de Atuação dos Examinadores por Subclasses da CIP	167
Tabela 98 – Medianas da Amostra de Redistribuição separadas por examinador.....	167
Tabela 99 – ZAE dos examinadores após a nova distribuição	168
Tabela 100 – Porcentagem de pedidos dentro das ZAE antes e após a nova distribuição	168
Tabela 101 – Medianas da Amostra de Redistribuição separadas por examinador.....	169
Tabela 102 – Resultados de p-valor dos Testes de Mann-Whitney	170

LISTA DE FIGURAS

Figura 1 – Exemplo de Folha de Rosto de Pedido de Patente de Invenção.....	24
Figura 2 – Distribuição normal adaptada do portal action.	49
Figura 3 – Diagrama do Modelo	50
Figura 4 – Exemplo de Quadro 1 do parecer de exame do INPI.....	53
Figura 5 – Exemplo de Quadro 2 do parecer de exame do INPI.....	53
Figura 6 – Exemplo de Quadro 3 do parecer de exame do INPI.....	55
Figura 7 – Exemplo de Quadro 4 do parecer de exame do INPI.....	55
Figura 8 – Exemplo de Quadro 5 do parecer de exame do INPI.....	55
Figura 9 – Formulário de Tempo de Exame (Parte 1).....	56
Figura 10 – Formulário de Tempo de Exame (Parte 2).....	56
Figura 11 – IGC padronizado dos pedidos.....	72
Figura 12 – Enquadramentos dos pedidos da amostra de teste inicial	73
Figura 13 – Caso 9 Var: autovalores para amostras de 10 até 100 pedidos	86
Figura 14 – Caso 9 Var: variância acumulada para amostras de 10 até 100 pedidos	86
Figura 15 – Caso 9 Var: correlações do IGC com o tempo.....	88
Figura 16 – Caso 9 Var: correlações do IGC_{Y1} , número de páginas de reivindicações e número total de páginas com o tempo.....	89
Figura 17 – Caso 5 Var: autovalores para amostras de 10 até 100 pedidos	91
Figura 18 – Caso 5 Var: variância acumulada para amostras de 10 até 100 pedidos.....	91
Figura 19 – Caso 5 Var: correlações do IGC_{Y1} , IGC, número de páginas de reivindicações e número total de páginas com o tempo.....	93
Figura 20 – Caso 4 Var: autovalores para amostras de 10 até 100 pedidos	95
Figura 21 – Caso 4 Var: variância acumulada para amostras de 10 até 100 pedidos.....	96
Figura 22 – Caso 4 Var: correlações do IGC_{Y1} , número de páginas de reivindicações e número total de páginas com o tempo.....	98
Figura 23 – Caso 3 Var: autovalores para amostras de 10 até 100 pedidos	100
Figura 24 – Caso 3 Var: variância acumulada para amostras de 10 até 100 pedidos.....	100
Figura 25 – Caso 3 Var: correlações do IGC com o tempo.....	102
Figura 26 – Caso 3 Var (2): autovalores para amostras de 10 até 100 pedidos.....	104
Figura 27 – Caso 3 Var (2): variância acumulada para amostras de 10 até 100 pedidos	104
Figura 28 – Caso 3 Var (2): correlações do IGC com o tempo	106
Figura 29 – Correlações do IGC e IGC_{Y1} com o tempo para amostras de 10 pedidos.....	107
Figura 30 – Correlações do IGC e IGC_{Y1} com o tempo para amostras de 20 pedidos.....	107
Figura 31 – Correlações do IGC e IGC_{Y1} com o tempo para amostras de 30 pedidos.....	108
Figura 32 – Correlações do IGC e IGC_{Y1} com o tempo para amostras de 40 pedidos.....	108
Figura 33 – Correlações do IGC e IGC_{Y1} com o tempo para amostras de 50 pedidos.....	108
Figura 34 – Correlações do IGC e IGC_{Y1} com o tempo para amostras de 60 pedidos.....	109
Figura 35 – Correlações do IGC e IGC_{Y1} com o tempo para amostras de 70 pedidos.....	109
Figura 36 – Correlações do IGC e IGC_{Y1} com o tempo para amostras de 80 pedidos.....	109
Figura 37 – Correlações do IGC e IGC_{Y1} com o tempo para amostras de 90 pedidos.....	110
Figura 38 – Correlações do IGC e IGC_{Y1} com o tempo para amostras de 100 pedidos.....	110
Figura 39 – Correlações do IGC_{Y1} com o tempo para todos os 5 casos e amostras de 10 a 100 pedidos	111
Figura 40 – Tempo padronizado dos pedidos	113
Figura 41 – Enquadramentos por tempo dos pedidos da amostra padrão	113
Figura 42 – IGC_{Y1} padronizado dos pedidos.....	116
Figura 43 – Enquadramentos por IGC_{Y1} dos pedidos da amostra padrão	116
Figura 44 – Correlações do IGC_{Y1} com o tempo para todos os 5 casos e amostras de 10 a 100 pedidos	120
Figura 45 – IGC_{Y1} padronizado dos pedidos.....	124
Figura 46 – Enquadramentos por IGC_{Y1} dos pedidos da amostra padrão	124
Figura 47 – Caso 3 Var (2): autovalores para amostras de 10 até 50 pedidos.....	134
Figura 48 – Caso 3 Var (2): variância acumulada para amostras de 10 até 50 pedidos	134
Figura 49 – IGC_{Y1} padronizado dos pedidos.....	136
Figura 50 – Enquadramentos por IGC_{Y1} dos pedidos da amostra	136
Figura 51 – Caso 3 Var (2): autovalores para amostras de 10 até 50 pedidos.....	144
Figura 52 – Caso 3 Var (2): variância acumulada para amostras de 10 até 50 pedidos	144
Figura 53 – IGC_{Y1} padronizado dos pedidos.....	145
Figura 54 – Enquadramentos por IGC_{Y1} dos pedidos da amostra	146

Figura 55 – Caso 3 Var (2): autovalores para amostras de 10 até 50 pedidos.....	153
Figura 56 – Caso 3 Var (2): variância acumulada para amostras de 10 até 50 pedidos	154
Figura 57 – IGC _{Y1} padronizado dos pedidos.....	155
Figura 58 – Enquadramentos por IGC _{Y1} dos pedidos da amostra	155
Figura 59 – Caso 3 Var (2): autovalores para amostras de 10 até 50 pedidos.....	162
Figura 60 – Caso 3 Var (2): variância acumulada para amostras de 10 até 50 pedidos	163
Figura 61 – IGC _{Y1} padronizado dos pedidos.....	164
Figura 62 – Enquadramentos por IGC _{Y1} dos pedidos da amostra	164
Figura 63 – Comportamento das variâncias acumuladas das áreas analisadas.....	171
Figura 64 – Variâncias da componente principal Y1 das áreas analisadas	172
Figura 65 – Correlações do IGC _{Y1} com o número de páginas de reivindicações.....	173

LISTA DE SIGLAS

ACP – Análise de Componentes Principais

CIP – Classificação Internacional de Patentes

DIRPA – Diretoria de Patentes, Programas de Computador e Topografias de Circuitos Integrados

EPO – European Patent Office

HO – Home Office

IBD – Índice de Balanceamento da Distribuição

IGC – Índice Geral de Complexidade

IN – Instrução Normativa

INID – International Agreed Numbers for the Identification Data

INPI – Instituto Nacional da Propriedade Industrial

LPI – Lei de Propriedade Industrial

OMPI – Organização Mundial da Propriedade Intelectual

SISCAP – Sistema de Cadastramento da Produção

USPTO – United States Patent and Trademark Office

VBA – Visual Basic for Applications

WIPO – World Intellectual Property Organization

ZAE – Zona de Atuação Específica do Examinador

1 INTRODUÇÃO

A concessão de ativos de propriedade industrial (PI) deve se nortear por dois princípios centrais e inter-relacionados: qualidade e eficiência. Um sistema eficiente para a proteção dos direitos de PI é vinculado principalmente ao tempo do trâmite processual para a concessão, mas também à clareza e organização do processo de análise realizado no exame técnico. Já a qualidade é normalmente vinculada à padronização e aperfeiçoamento de procedimentos administrativos, proporcionando confiabilidade e segurança jurídica nas decisões.

Quanto ao trâmite processual e tempo de concessão dos direitos de propriedade industrial, o Instituto Nacional da Propriedade Industrial (INPI) vem promovendo esforços para que o tempo de concessão do Instituto seja reduzido e enquadrado na média dos principais escritórios internacionais. Dentre suas ações, destaca-se a implementação de um “plano de ataque ao backlog” anunciado em 2019, plano este já em vigor e com previsão de duração até 2021 (INPI, 2019). Cabe destacar os resultados e esforços quanto ao tempo de obtenção de uma marca pelo Instituto, culminando na inclusão do INPI entre os signatários do Protocolo de Madri (INPI, 2019).

Ao longo das últimas décadas, o INPI vem incorporando em seu Planejamento Estratégico iniciativas relacionadas à qualidade dos seus processos. No Plano Estratégico 2018-2021 (INPI, 2020) foram mencionadas parcerias, acordos de cooperação com escritórios internacionais, aperfeiçoamento dos programas de exame prioritário, elaboração e revisão de diretrizes de exame de patentes, dentre outras atividades. Quanto às iniciativas de cooperação técnica, destaca-se recente parceria estratégica entre o INPI e o Escritório de Patentes Europeu (EPO, do inglês *European Patent Office*) incluindo treinamentos, discussão de melhores práticas, compartilhamento de ferramentas e intercâmbio de bases de dados de patentes (INPI, 2019).

Ressalte-se ainda que, em 2018, por meio do Plano de Ação da Diretoria de Patentes do INPI (DIRPA), foi definido o objetivo estratégico 1 – otimizar a qualidade e o tempo para concessão de patentes – incluindo como iniciativa estratégica o item 1.5.1 – Projeto “Equilíbrio da carga de Exame”. Entre as principais ações relacionadas ao dito projeto, destacam-se:

“- Avaliar o perfil de classificação e de progressão de exame. Propor um balanceamento dentro das Coordenações. Detectar anomalias (nichos) tecnológicos – propor treinamento capacitação nestas áreas;
- Fazer levantamento de pedidos pendentes conforme área técnica detalhada;
- Propor esquema de redistribuição em lote;
- Reduzir a faixa de símbolos de classificações da IPC para cada examinador possibilitando a escolha de sub-áreas por examinador;
- Implementar no SISCAP recursos que possibilitem um melhor gerenciamento para distribuição de pedidos por classificação”.
(DIRPA, Plano de Ação, 2018).

Destaque-se que existem estudos como os de Saint-Georges e Potterie (2013), Philipp (2006) e Kim e Oh (2017) mostrando que variações na carga de trabalho de examinadores afetam a qualidade do processo de exame e seus resultados, sugerindo que um fator com elevado potencial de geração de instabilidade nos processos e divergências nas decisões de um escritório de patentes é o desbalanceamento de carga dos examinadores de patente, ou seja, uma distribuição de pedidos de patente desequilibrada entre eles. Adicionalmente, o trabalho de Golodne e Silva (2020), um estudo específico utilizando dados de pedidos de patente de diversas áreas do INPI, mostra que o volume de dados dos pedidos de patente pode ser usado para avaliar a carga de trabalho dos examinadores e sugere que páginas de reivindicações seria uma variável chave para esta avaliação. Tal fato torna-se ainda mais relevante uma vez que não existe uma metodologia de distribuição formal atualmente no INPI, sendo que a decisão/responsabilidade da distribuição fica a cargo dos ocupantes de cargos de chefia de divisão, cada um com a liberdade de aplicar seus próprios critérios.

Assim, devido à relevância do tema “equilíbrio da carga de exame” para o próprio INPI, especificamente no que tange à DIRPA, e à ausência, na literatura, de estudos sobre metodologias de análise e distribuição de pedidos de patente aos examinadores em busca de proporcionar um balanceamento da carga otimizado, o desenvolvimento de uma metodologia se revela uma grande contribuição à eficiência e qualidade do exame de patentes, sendo um elemento importante para o alcance da missão do INPI e também para outros escritórios internacionais de Propriedade Intelectual.

1.1 OBJETIVOS

O objetivo geral da tese foi desenvolver uma metodologia para distribuição dos pedidos de patente na carga dos examinadores que leve em consideração, além do

enquadramento do pedido na Classificação Internacional de Patentes (CIP), variáveis relacionadas ao volume de dados do pedido e sua complexidade para o processo de exame.

Tendo definido qual é o objetivo geral da tese, são então definidos os objetivos específicos:

1. Identificar as variáveis relevantes do pedido de patente para realização de uma avaliação e distribuição justa e eficiente;
2. Identificar as Zonas de Atuação Específicas (ZAE) de cada examinador a ser avaliado, ou seja, determinar quais as subclasses da Classificação Internacional de Patentes (CIP) que examinam com maior frequência;
3. Criar um modelo com índices para avaliação de cada um dos pedidos de patente de acordo com as variáveis identificadas, por meio de uma análise multivariada;
4. Analisar a acurácia do modelo criado avaliando sua correlação com o tempo de exame;
5. Propor uma nova metodologia de distribuição de pedidos de patente baseada nas ZAE dos examinadores e no modelo criado;
6. Criar um Índice de Balanceamento para comparação das distribuições e redistribuições realizadas.

1.2 ORGANIZAÇÃO DO TEXTO

A apresentação da tese está dividida em seis seções, incluindo esta introdução. No Capítulo 2 são apresentados a estrutura básica de um pedido de patente e da Classificação Internacional de Patentes (CIP), além de estudos diversos utilizando como base variáveis do pedido de patente. Neste contexto, são identificadas as possíveis principais variáveis de um pedido de patente com impacto na sua distribuição aos examinadores. Já no Capítulo 3 é apresentada a metodologia utilizada para levantamento dos dados e execução de cada uma das etapas de pesquisa e análise para cumprimento dos objetivos previamente definidos.

O Capítulo 4 apresenta o desenvolvimento executado e a base teórica no qual é fundamentado. Ele é dividido em três partes principais. Na primeira, é realizada uma análise detalhada da técnica multivariada utilizada, com apresentação do passo a passo para sua execução, incluindo cálculos, critérios adotados e testes para verificação de adequação dos

dados. Adicionalmente, é apresentada a metodologia desenvolvida para avaliação dos pedidos de patente compreendendo um Índice Geral de Complexidade (IGC) dos pedidos de patente e uma forma de enquadramento dos pedidos em até cinco classes de complexidade, apresentando assim o modelo de avaliação criado. Já na segunda parte, o exame substantivo de patentes e o parecer padrão de exame do INPI são analisados, sendo construído um formulário de tempo de exame baseado nas principais etapas de exame, servindo de base para obtenção da Amostra Padrão de pedidos de patente a ser usada para realização de simulações específicas visando a uma análise de sensibilidade do modelo. Por fim, a terceira parte apresenta uma lógica para distribuição dos pedidos de patente avaliados pelo modelo, além de um Índice de Balanceamento para realizar uma comparação entre a distribuição real e a proposta, utilizando como base a Amostra Final de Redistribuição obtida.

Em seguida apresentam-se, no Capítulo 5, os resultados obtidos e discussões. Já no capítulo 6 são descritas as conclusões e as contribuições para futuras pesquisas.

2 O DOCUMENTO DE PATENTE E AS VARIÁVEIS DE INTERESSE PARA AVALIAÇÃO E DISTRIBUIÇÃO

2.1 ESTRUTURA BÁSICA DO PEDIDO DE PATENTE

A Lei de Propriedade Industrial (LPI), em seu capítulo III – artigo 19, define que:

“Art. 19. O pedido de patente, nas condições estabelecidas pelo INPI, conterá:

- I - requerimento;
- II - relatório descritivo;
- III - reivindicações;
- IV - desenhos, se for o caso;
- V - resumo; e
- VI - comprovante de pagamento da retribuição relativa ao depósito.”

Portanto, um pedido de patente é composto basicamente por quatro partes: relatório descritivo, reivindicações/quadro reivindicatório, desenhos (quando necessário) e resumo. Destaque-se que estas quatro partes principais apresentam características específicas de forma e conteúdo que serão avaliadas em maior detalhe.

2.1.1 Relatório Descritivo

Com relação ao relatório descritivo o artigo 24 da LPI rege que:

“Art. 24. O relatório deverá descrever clara e suficientemente o objeto, de modo a possibilitar sua realização por técnico no assunto e indicar, quando for o caso, a melhor forma de execução.”

Adicionalmente, cabe destacar que a Instrução Normativa (IN) Nº 030/2013 (IN 030/2013) editada pelo INPI, ainda menciona que:

- “Art. 2º O relatório descritivo deverá cumprir as seguintes especificações:
- V. definir os objetivos da invenção e descrever, de forma clara, concisa e precisa, a solução proposta para o problema existente, bem como as vantagens da invenção em relação ao estado da técnica;
 - VI. ressaltar, nitidamente, a novidade e evidenciar o efeito técnico alcançado;
 - VII. relacionar as figuras apresentadas nos desenhos, especificando suas representações gráficas (vistas, cortes, [...]);
 - VIII. descrever a invenção de forma consistente, precisa, clara e suficiente, fazendo remissão aos sinais de referência constantes dos desenhos, se houver, [...];

X. Indicar, explicitamente, a utilização industrial quando esta não for evidente a partir da descrição da invenção;”.

De todo o exposto, verifica-se que o relatório descritivo é o local onde o requerente/inventor deve descrever em detalhes a sua invenção, deixando claro para terceiros e para um técnico no assunto como se daria a sua execução na prática. Consequentemente, é uma das partes de maior relevância do pedido de patente, e é de se esperar que seja uma das partes de maior extensão (número de páginas). Portanto, o número de páginas do relatório descritivo de um pedido de patente é uma variável relevante para sua avaliação.

2.1.2 Reivindicações

Quanto às reivindicações o artigo 25 da LPI rege que:

“Art. 25. As reivindicações deverão ser fundamentadas no relatório descritivo, caracterizando as particularidades do pedido e definindo, de modo claro e preciso, a matéria objeto de proteção.”

Ressalte-se que, em um mesmo quadro reivindicatório, as reivindicações podem ser classificadas como independentes ou dependentes, conforme descrito na IN 030/2013:

“Art. 5º As reivindicações independentes deverão cumprir as seguintes especificações:

I. as reivindicações independentes visam a proteção de características técnicas essenciais e específicas da invenção em seu conceito integral, cabendo a cada categoria de reivindicação pelo menos uma reivindicação independente;

V. após a expressão “caracterizado por” devem ser definidas as características técnicas essenciais e particulares que, em combinação com os aspectos explicitados no preâmbulo, se deseja proteger;

“Art. 6º As reivindicações dependentes deverão cumprir as seguintes especificações:

I. são aquelas que, mantida a unidade de invenção, incluem todas as características de outra(s) reivindicação(ões) anterior(es) e definem detalhamentos dessas características e/ou características adicionais que não sejam consideradas características essenciais da invenção, devendo conter uma relação de dependência a essa(s) reivindicação(ões) e a expressão “caracterizado por”;

Deste modo, no bojo de tudo que foi exposto em grande extensão ao longo do relatório descritivo, as reivindicações ou o quadro reivindicatório especificam de maneira mais exata o objeto de proteção, ou seja, definem o escopo de proteção. Em estudo recente de Soares, Moreira Jr, Rodrigues e Braga (2019) o papel do termo caracterizante (expressão

“caracterizado por”) nos documentos de patente na delimitação de direitos de propriedade industrial é discutido em mais detalhes considerando o histórico do arcabouço legal brasileiro e internacional.

Neste contexto, nota-se que tanto a extensão das reivindicações (retratada pelo seu número de páginas) quanto o número de reivindicações (soma do número de reivindicações independentes e dependentes) são de grande relevância quando se pretende avaliar um pedido de patente.

2.1.3 Desenhos

De acordo com o disposto no artigo 19 da LPI, é facultativa a apresentação de desenhos no pedido de patente. Todavia, em caso de apresentação, eles devem seguir o disposto na IN 030/2013:

“Art. 8º Os desenhos, fluxogramas e diagramas, esquemas gráficos deverão:
I. ser isentos de textos, rubricas ou timbres, podendo conter apenas termos indicativos e palavras-chave;
II. conter todos os sinais de referência, tais como algarismos, letras ou alfanuméricos, constantes do relatório descritivo, observando o uso dos mesmos sinais de referência para identificar determinada característica em todos os desenhos, sempre que essa apareça;”

Embora os desenhos não sejam obrigatórios, nota-se que os requerentes e inventores, na sua maioria, os inserem na documentação apresentada para análise. Tal fato evidencia que os desenhos são parte importante do pedido de patente e, umas das principais razões para isso, é que eles facilitam a leitura e o entendimento da matéria em análise. A relevância dos desenhos torna-se ainda mais evidente quando da análise dos seguintes dispositivos da IN 030/2013 que correlacionam os desenhos tanto com o relatório descritivo quanto com as reivindicações:

“Art. 2º O relatório descritivo deverá cumprir as seguintes especificações:
VII. relacionar as figuras apresentadas nos desenhos, especificando suas representações gráficas (vistas, cortes, [...]);
VIII. descrever a invenção de forma consistente, precisa, clara e suficiente, fazendo remissão aos sinais de referência constantes dos desenhos, se houver, [...];”

“Art. 4º As reivindicações, quanto à formulação, deverão cumprir as seguintes especificações:
VI. quando o pedido contiver desenhos, as características técnicas definidas nas reivindicações devem vir acompanhadas, entre parênteses, pelos respectivos sinais de referência constantes dos desenhos se for considerado necessário à compreensão do mesmo, entendendo que tais sinais de referência não são limitativos das reivindicações;”.

Vale ressaltar ainda que boa parte dos exames técnicos apresentam exigências que apontam justamente a ausência de referências numéricas de um pedido de patente e consequente prejuízo de sua clareza por conta disso. Sendo assim, verifica-se que os desenhos, embora não apresentando a mesma relevância do relatório descritivo e do quadro reivindicatório, acabam por complementá-los e, portanto, também são relevantes quando se pretende avaliar um pedido de patente.

2.1.4 Resumo

Quanto ao Resumo, a IN 030/2013 faz as seguintes considerações:

“Art. 7º O resumo, quanto ao conteúdo, deverá cumprir as seguintes especificações:

I. ser iniciado pelo título e ser um sumário do que foi exposto no relatório descritivo;

II. indicar o setor técnico ao qual pertence a invenção;

III. ser redigido de forma a permitir uma compreensão clara do problema técnico, da essência da solução desse problema por meio da invenção e do uso principal ou dos usos principais da invenção;

IV. ser redigido de forma a poder servir de instrumento eficaz de pré-seleção para fins de pesquisa em determinado setor técnico, especialmente ajudando o usuário a formular uma opinião quanto à conveniência ou não de consultar o documento na íntegra.”

Adicionalmente, cabe destacar que a Instrução Normativa Nº 031/2013 (IN031/2013) editada pelo INPI, ainda faz as seguintes considerações acerca do resumo de um pedido de patente:

“Art. 22 O resumo, quanto à forma, deverá cumprir as seguintes especificações:

I. ser iniciado pelo título, em destaque com relação ao restante do texto;

II. indicar o setor técnico ao qual pertence a invenção;

III. ser tão conciso quanto a exposição permitir (de preferência de 50 a 200 palavras), não excedendo 25 linhas de texto.”

Verifica-se que o resumo apresenta alguma relevância principalmente no que se refere ao inciso IV do artigo 7º, ou seja, serve para uma verificação inicial da conveniência de se consultar ou não o documento na íntegra, seja para uma análise específica de interesse do usuário, seja para uma busca inicial de estado da técnica por parte do examinador de patentes.

Todavia, conforme explicitado no inciso III do artigo 27 da IN 031/2013, o resumo deve ser bem conciso, não excedendo 25 linhas, ou seja, acaba sendo limitado a uma única folha, ainda assim não totalmente preenchida. Por conseguinte, ainda que se reconheça a importância do resumo, nota-se que sua influência é bastante reduzida no que tange ao

volume de dados e/ou complexidade do pedido de patente e, sendo assim, não será considerado em análises posteriores.

2.2 FOLHA DE ROSTO: DADOS BIBLIOGRÁFICOS DO DOCUMENTO DE PATENTE

Oliveira, Suster, Pinto, Ribeiro e Silva (2005) mostram que informação contida nos documentos de patente, aí incluídos os dados bibliográficos constantes da folha de rosto, é uma ferramenta importante para a pesquisa e o desenvolvimento tecnológico. Vale ressaltar que as informações contidas na folha de rosto de um documento de patente são identificadas por códigos numéricos, mais conhecidos como códigos INID (*International Agreed Numbers for the Identification Data*), sendo que seus padrões específicos são definidos pela Organização Mundial da Propriedade Intelectual (OMPI) ou, do inglês, *World Intellectual Property Organization* (WIPO). A Figura 1 mostra um exemplo de folha de rosto de um documento de patente.

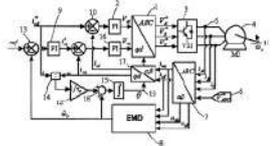
 República Federativa do Brasil Ministério do Desenvolvimento, Indústria e do Comércio Exterior Instituto Nacional da Propriedade Industrial	<p>(21) PI 1105811-0 A2</p> <p>(22) Data do Depósito: 04/11/2011</p> <p>(43) Data da Publicação: 05/01/2016 (RPI 2348)</p>	 * BR P 2 1 1 0 5 8 1 1 A 2 *
<hr/>		
<p>(54) Título: SISTEMA DE CONTROLE DE VELOCIDADE ROTÓRICA E MÉTODO DE ESTIMAÇÃO DA VELOCIDADE ROTÓRICA</p> <p>(51) Int. Cl.: H02P 21/14; H02P 21/12</p> <p>(73) Titular(es): UNIVERSIDADE FEDERAL DE SANTA MARIA</p> <p>(72) Inventor(es): HILTON ABÍLIO GRÜNDLING, RODRIGO PADILHA VIEIRA</p>	<p>(57) Resumo: SISTEMA DE CONTROLE DE VELOCIDADE ROTÓRICA E MÉTODO DE ESTIMAÇÃO DA VELOCIDADE ROTÓRICA. A presente invenção compreende um sistema para controle da velocidade rotórica de máquinas de indução, compreendendo módulo de acionamento, interface de processamento digital de sinais e máquina de indução. A presente invenção compreende, adicionalmente, um método de estimação de velocidade rotórica, sem o uso de sensores mecânicos, por modos deslizantes em tempo discreto, o que permite que a invenção possa ser utilizada diretamente em processadores digitais de sinais ou microcontroladores sem a necessidade da discretização de equações diferenciais. O sistema de controle da velocidade rotórica pode utilizar qualquer tipo de lei de controle capaz de ser implementada em um processador digital de sinais desde que a mesma seja projetada para ser estável, e a informação da velocidade atual do eixo do rotor é obtida a partir de informações das variáveis elétricas do motor, as quais são utilizadas por um algoritmo de estimação de velocidade.</p>	
		

Figura 1 – Exemplo de Folha de Rosto de Pedido de Patente de Invenção

Nota-se que os dados bibliográficos contidos na folha de rosto apresentam informações importantes relativas aos seguintes eixos:

- ✓ Números específicos para controle e padronização, tais como:
 - (21) Número designado ao documento quando do depósito;
 - (31) Número designado ao primeiro depósito/Prioridade, se houver.
- ✓ Marcos temporais, tais como:
 - (22) Data de depósito;
 - (32) Data do primeiro depósito/Prioridade, se houver;
 - (43) Data de Publicação;
- ✓ Dados de Identificação propriamente dita, tais como:
 - (72) Nome do inventor, se conhecido;
 - (73) Nome do detentor de direitos sobre a patente;
 - (75) Nome do inventor, quando também for o depositante.
- ✓ Dados com Indicação Técnica, tais como:
 - (51) Classificação Internacional de Patente;
 - (54) Título da Invenção;
 - (57) Resumo do conteúdo do documento.

Uma descrição detalhada destes códigos pode ser consultada no apêndice 1 da Norma ST.9 do Manual de Informação e Documentação em Matéria de Propriedade Intelectual da OMPI (2013).

Neste contexto cabe destacar que, como se pretende obter variáveis que sejam relevantes para a avaliação e distribuição dos pedidos de patente para exame, destes dados bibliográficos, só irão nos interessar aqueles que já estão disponíveis desde o depósito até a publicação do pedido, ou seja, aqueles que podem ser obtidos antes do processo de exame de patente no INPI. Sendo assim, em uma primeira análise há interesse nos números de identificação dos pedidos, data/ano de depósito, inventores, prioridades e a CIP, sendo que esta última será abordada em maiores detalhes a seguir.

2.3 CLASSIFICAÇÃO INTERNACIONAL DE PATENTES (CIP)

A CIP foi o meio encontrado para uniformizar os documentos de diversos países com linguagens e expressões técnicas diferentes. De acordo com a OMPI/WIPO, tem papel importante no sentido de ser: i) uma ferramenta para arranjar ordenadamente os documentos de patentes, de modo a auxiliar o acesso às informações técnicas e legais nele apresentadas; ii) uma base de divulgação seletiva de informações a todos que utilizam patentes como referência e/ou conhecimento; iii) um suporte para fazer um levantamento do estado da arte referente a alguns campos da tecnologia; e iv) um apoio para desenvolvimento de estatísticas a respeito da propriedade industrial, permitindo analisar o avanço tecnológico em diversas áreas (DELVIZIO, 2010).

A CIP está dividida em oito seções (A, B, C, D, E, F, G e H). As seções são os níveis mais altos da hierarquia da classificação. Cada seção é subdividida em classes, que representam o segundo nível hierárquico da classificação. Cada classe abrange uma ou mais subclasses, que indicam o terceiro nível hierárquico da classificação. Cada subclasse é desdobrada em subdivisões, denominadas “grupos”, que são tanto grupos principais (quarto nível hierárquico da classificação) quanto subgrupos (níveis hierárquicos mais baixos dependentes do nível do grupo principal da classificação). A seguir tem-se um exemplo de uma classificação da área de eletricidade utilizando todos os níveis hierárquicos, desde a seção até o subgrupo.

Exemplo: **H02M 7/48**

Onde:

H Seção (A, B, [...], H): Eletricidade;

H02 Classe (dois dígitos): Produção, Conversão ou Distribuição de Energia;

H02M Subclasse (uma letra): Aparelhos para Conversão CC-CA;

H02M7 Grupo Principal (um ou mais dígitos numéricos): Conversão de uma potência de entrada CC em uma potência de saída CA;

H02M7/48 Subgrupo – (pelo menos dois dígitos numéricos): Por meio de conversores usando semicondutores com eletrodo de controle.

Destaque-se que a CIP é periodicamente revisada a partir de reuniões de peritos dos países membros da OMPI/WIPO, sendo divulgada e podendo ser acessada por meio dos *sites* do INPI e da OMPI/WIPO.

Cumprido destacar ainda que um documento de patente pode ser enquadrado em mais de uma CIP. Isto se deve ao fato de poder apresentar reivindicações de mais de uma categoria

(dispositivo e método, por exemplo), estar relacionada a um ou mais campos de aplicação ou mesmo a funções específicas.

No estudo de Harloff e Wagner (2006), por meio de dados do EPO, foi realizado um trabalho para modelagem da duração do tempo de exame em que definiu que, dentre outros fatores que impactam na complexidade de um pedido de patente, o número total de reivindicações e o número de classificações da CIP são variáveis diretas do pedido de grande relevância. Diante disso, o número de subclasses da CIP também é uma variável a ser considerada para avaliação de um pedido de patente a ser distribuído.

2.4 SUBSÍDIOS AO EXAME TÉCNICO

Percebe-se que a estrutura do pedido de patente, a folha de rosto do documento de patente e a estrutura da CIP permitem obter uma gama de variáveis relevantes a serem consideradas quando da avaliação de um pedido de patente. Todavia, um importante elemento a ser considerado refere-se à possibilidade de apresentação de subsídios ao exame técnico enviados por terceiros interessados. A LPI, em seu artigo 31, define que:

“Art. 31. Publicado o pedido de patente e até o final do exame, será facultada a apresentação, pelos interessados, de documentos e informações para subsidiarem o exame.”

Adicionalmente, cabe destacar que a Norma de Execução Nº 4/2015 editada pelo INPI institui o Procedimento de Avaliação de Subsídios de Terceiros ao Exame Técnico (DIRPA - P026, 2015) que ainda menciona:

“5. PROCEDIMENTO

5.1 Citar a petição de subsídio no parecer técnico, mencionando o número, a data e o nome do interessado, usando a seguinte redação: “Todos os documentos apresentados no subsídio através da petição [MMM] de XX/YY/ZZ depositada por [NN] foram avaliados de acordo com o artigo 31 da lei 9279/96”.

5.2 Avaliar a documentação apresentada para fins de verificação do cumprimento dos requisitos de patenteabilidade.”.

Verifica-se que a LPI faculta a terceiros interessados a apresentação de subsídios ao exame técnico, sendo que não há limitação de número de documentos nem páginas a serem apresentadas. Além disso, o procedimento de exame OMPI-P026 obriga o examinador a analisar toda a documentação apresentada e citar explicitamente no parecer técnico confirmando que a analisou. Sendo assim, o número de páginas apresentado como subsídios

pode ser bastante extenso e influi diretamente no trabalho de análise a ser executado pelo examinador de patentes. Consequentemente, o número de páginas de subsídios é, em princípio, uma variável de grande relevância para avaliação do pedido de patente a ser distribuído.

Por outro lado, vale destacar que a quantidade de pedidos de patente incluindo subsídios de terceiros depositados no INPI é baixa. Mais especificamente, na área de eletricidade, estes depósitos normalmente não passam em média de cinco a dez pedidos ao ano, ou seja, menos de um pedido por ano por examinador, quantidade que não justifica esforços maiores para um modelo de avaliação focado nesta variável.

2.5 ESTUDOS UTILIZANDO VARIÁVEIS DOS DOCUMENTOS DE PATENTE

2.5.1 Estudos diversos

Um pedido de patente é composto basicamente por quatro partes: relatório descritivo, quadro reivindicatório, desenhos (quando necessário) e resumo. No caso do relatório descritivo e dos desenhos, por exemplo, o número de páginas é uma variável relevante. Já com relação ao quadro reivindicatório, o número de páginas, número de reivindicações, se são independentes ou não, etc. Além disso, os dados bibliográficos apresentam informações que se tornam variáveis também possivelmente relevantes, tais como as classificações da CIP, inventores, datas e anos de depósito etc. Neste contexto, existem alguns trabalhos que utilizam variáveis específicas dos documentos de patente para determinar características como: valor econômico, esforço de tempo na avaliação técnica, escopo de proteção da patente, dentre outros.

Em Gambardella, Harhoff e Verspagen (2008) há uma abordagem acerca do valor econômico de patentes europeias utilizando como variáveis suas reivindicações, citações, dentre outros.

Já os trabalhos de Kuhn e Thompson (2017) e Marco, Sarnoff e deGrazia (2016) tratam de metodologias para análise e medição do escopo de proteção de uma patente utilizando como base variáveis relacionadas ao quadro reivindicatório. No primeiro é utilizado o número de palavras da primeira reivindicação independente da patente como variável relevante. No segundo, são utilizadas duas variáveis: número de palavras da menor reivindicação independente da patente e o número total de reivindicações independentes.

No entanto, nenhum dos trabalhos supramencionados trata da utilização dos parâmetros do pedido de patente ou de dados do processo de exame para criação de uma metodologia de distribuição dos pedidos para os examinadores. Em Cockburn, Kortum e Stern (2002) e Lemley e Sampat (2012), foram realizados estudos com evidências de que, no Escritório Americano de Patentes (USPTO), após o pedido ser direcionado a uma grande área tecnológica (por exemplo, eletricidade ou química), sua posterior designação a um examinador específico era realizada quase que de maneira aleatória.

No recente trabalho de Cesare e Timothy (2018) foram encontradas evidências de que, no USPTO, pedidos de um mesmo depositante, bem como pedidos com similaridades nos resumos e títulos, tendem a ser enviados para o mesmo examinador. Os autores sugerem que embora tais metodologias busquem seguir em consonância com o princípio da eficiência, há de se considerar também um balanço com o princípio de justiça. Por outro lado, ressaltam que uma distribuição aleatória privilegia o princípio de justiça, mas peca no atendimento ao princípio da eficiência, sobre o qual pesa a especialização dos examinadores em determinadas subclasses de exame.

Diante do exposto, para que haja eficiência e justiça, há de se considerar não apenas os fatores que contribuam com a eficiência, tais como a distribuição de pedidos por subclasses de acordo com a formação/interesse dos examinadores, como também fatores que contribuam com a justiça e a igualdade, ou seja, variáveis que retratem o volume de dados e a complexidade dos pedidos de patente a serem analisados. Por conseguinte, justiça não pode ser confundida com aleatoriedade.

2.5.2 Balanceamento de Carga, Voluminosidade e Complexidade dos Pedidos de Patente

Um estudo de Saint-Georges e Potterie (2012) mostra que a carga de trabalho de cada escritório de patentes é um dos indicadores que afetam a qualidade não somente do exame de patentes em si, mas dos sistemas patentários como um todo. Cabe destacar ainda os trabalhos de Philipp (2006) e Kim e Oh (2017) que indicam que um aumento da carga de trabalho dos examinadores, e consequente redução de tempo para análise de um pedido de patente, pode afetar negativamente a qualidade de patentes concedidas, ou seja, tende a aumentar a concessão de patentes indevidas. Sendo assim, torna-se importante a busca por uma melhor

distribuição e balanceamento de carga dos examinadores, de modo a obter uma harmonização das condições de trabalho e de resultados.

Os trabalhos de Archontopoulos, Guellec, Stevnsborg, de la Potterie, e Zeebroeck (2007, 2008) mostram que o volume dos pedidos de patente no EPO tem aumentado ao longo dos anos e associam o volume do pedido de patente a um problema essencialmente bidimensional, com predominância de duas variáveis: total de páginas do pedido de patente e número de reivindicações. Neste contexto o total de páginas estaria relacionado ao grau de descrição do invento fornecido e o total de reivindicações ao escopo ou abrangência de proteção. Os autores destacam também que estas variáveis estariam correlacionadas, em diferentes proporções, com o número de prioridades, inventores e classificações da CIP. Como limitações do trabalho citam ainda que seria vantajoso avaliar a quantidade de reivindicações independentes e dependentes, embora esta informação não esteja facilmente acessível nas fontes de dados disponíveis.

Para avaliar a hipótese de complexidade tecnológica de um pedido de patente, outro estudo de van Zeebroeck, de la Potterie, e Guellec (2009) utiliza um modelo matemático considerando o número de inventores, número de classes da CIP e número de citações a patentes anteriores. Os autores ressaltam que o aumento no volume de dados pode tanto ser relacionado à complexidade técnica da invenção e do pedido em si quanto a uma questão de estratégia, ou seja, de se querer manter consigo certo *know-how* do que de fato pretende proteger. Importante ressaltar que, com os resultados do modelo aplicado, os autores sugerem que um maior número de inventores tende a dar origem a invenções mais complexas, sendo necessário maior número de páginas e de reivindicações para que ditas invenções sejam bem descritas.

2.6 VARIÁVEIS SELECIONADAS E CONSIDERAÇÕES INICIAIS

Da análise da estrutura dos pedidos de patente e arcabouço legal relacionado, além da revisão bibliográfica de estudos diversos utilizando variáveis dos pedidos de patente, inicialmente foram identificadas doze possíveis variáveis candidatas a serem utilizadas no presente trabalho:

- ✓ número de páginas do relatório descritivo (variável 1);
- ✓ número de páginas de reivindicações (variável 2);
- ✓ número de páginas de desenhos (variável 3);
- ✓ número de páginas de subsídios (variável 4);

- ✓ número de reivindicações independentes (variável 5);
- ✓ número de reivindicações dependentes (variável 6);
- ✓ número de subclasses da CIP (variável 7);
- ✓ ano de depósito (variável 8);
- ✓ número de inventores (variável 9);
- ✓ número de prioridades (variável 10);
- ✓ número de citações do estado da técnica pelo documento de patente em análise (variável 11); e
- ✓ número de citações do documento de patente em análise por outros documentos de patente (variável 12).

Dos estudos realizados com dados do EPO acerca da voluminosidade (Volume de dados) de um pedido de patente, infere-se que o volume de dados com que o examinador precisa lidar ao longo do exame de um pedido de patente é uma das principais condicionantes do esforço/tempo de exame. Além disso, tais estudos indicam que, de maneira simplificada, a voluminosidade é um problema relacionado a duas variáveis do pedido de patente: número total de páginas e número total de reivindicações. Sendo assim, uma primeira premissa considerará que o volume de dados de um pedido de patente pode ser representado pelas variáveis de maior correlação positiva com o número total de páginas do pedido e/ou com o número total de reivindicações.

Por outro lado, embora boa parte do esforço de exame esteja ligada ao volume direto de dados que o examinador lida, observa-se que há ainda a hipótese de existência de variáveis complementares, ligadas a uma complexidade indireta e mais subjetiva, especificamente relacionada ao pedido de patente em si, à estratégia dos requerentes ou mesmo a particularidades do processo de exame. Neste contexto, verifica-se que há algumas variáveis sugeridas nos estudos avaliados, porém, não há consenso dos autores pesquisados nem estudos que revelem a exata influência de cada uma delas ou não.

Por conseguinte, construímos duas hipóteses iniciais:

- ✓ Hipótese 1 - são cinco as variáveis diretamente ligadas ao volume de dados: número de páginas de relatório descritivo, reivindicações e desenhos, além do número de reivindicações dependentes e independentes.
- ✓ Hipótese 2 - ainda que reduzida, há influência indireta das demais variáveis identificadas.

3 METODOLOGIA

A metodologia proposta visa a construir um modelo capaz de avaliar a complexidade e volume dos pedidos de patente, além de uma nova lógica justa e eficiente para distribuição de pedidos de patente aos examinadores do INPI. Para tanto, é necessário identificar e selecionar as variáveis do pedido de patente que podem influenciar no exame, avaliar os pedidos de patente conforme as variáveis selecionadas, criar uma lógica específica para esta distribuição e, por fim, avaliar a nova lógica proposta comparando com a distribuição original. Neste contexto, a metodologia proposta pode ser dividida em cinco partes principais, que se complementam entre si e serão detalhadas a seguir.

3.1 IDENTIFICAÇÃO DAS VARIÁVEIS DE INTERESSE

Esta 1ª parte consistiu em avaliar a estrutura básica de um pedido de patente com a análise de documentos reais (Documentos apresentados nas petições de depósito) consultados por meio do sistema de cadastramento de produção do INPI e revisão do arcabouço legal específico (leis, diretrizes, procedimentos, instruções normativas, guias e artigos). Foi também realizada uma análise dos dados bibliográficos de um documento de patente por meio de pesquisa e consulta a documentos reais, além de uma revisão e análise dos principais estudos que utilizam como base as variáveis de documentos de patente.

Destaque-se que a formatação da análise documental com base nos documentos de patente envolve a realização das seguintes atividades:

- Definição da área tecnológica utilizada como referência para obtenção dos dados;
- Seleção do período temporal para análise;
- Elaboração de arquivos de dados contendo os resultados encontrados;
- Análise das informações encontradas.

Nesta etapa foi definido que a avaliação seria para a área de eletricidade da DIRPA. Adicionalmente, as premissas iniciais para estipular a janela temporal foram as seguintes:

- Definir um período que garantisse a publicação dos pedidos de patente, conforme Artigo 30 da LPI;
- Priorizar um período que já estivesse utilizando a última Instrução Normativa publicada pelo INPI (IN n^o 030/2013) e o novo Parecer Padrão do INPI;

- Definir um período em que o maior número possível de examinadores a serem avaliados já estivesse:
 - Estável; ou
 - Realizando seus exames/atividades sem supervisão/tutoria.

Neste contexto, para a realização do presente estudo, foram obtidos dados de todos os pedidos com primeiros exames já realizados ao longo de dois anos da área de eletricidade, mais precisamente de maio de 2015 até maio de 2017, mês este em que se deu início à pesquisa.

Com as possíveis variáveis de interesse identificadas, foram estabelecidos os seguintes critérios adicionais para seleção das variáveis do estudo:

- momento de disponibilidade da variável: como o objetivo principal é avaliar o documento para posterior distribuição para exame, serão selecionadas apenas variáveis disponíveis para acesso desde o depósito até a publicação do pedido de patente;
- confiabilidade e eficiência: não serão selecionadas variáveis sem que haja total confiança nos dados obtidos, que possam exigir a utilização de mais de uma plataforma/base de dados e/ou que exijam um tempo de obtenção muito elevado, o que poderia aumentar demais a complexidade do modelo ou mesmo inviabilizar a sua aplicação prática.

Com a identificação e seleção das variáveis, a proposta foi realizar um tabelamento dos dados de todos os pedidos de patente consultados, incluindo todas as variáveis relevantes selecionadas, além da identificação do examinador de patentes que recebeu o pedido para análise. Estes dados foram definidos como sendo a Amostra Inicial de Teste e, a partir dela, foram então identificadas as CIP dos pedidos de patente da amostra e, conseqüentemente, as ZAE de cada um dos examinadores.

3.2 AVALIAÇÃO DOS PEDIDOS DE PATENTE QUANTO AO VOLUME DE DADOS E COMPLEXIDADE

Esta 2ª parte consistiu na identificação e implementação de um método para avaliação dos pedidos de patente levando em consideração as variáveis relevantes selecionadas na 1ª parte. Em se tratando de um problema com múltiplas variáveis de interesse, buscou-se um

método de análise multivariada para solução. A revisão bibliográfica foi realizada com objetivo de selecionar um método que atendesse aos seguintes critérios:

- Não apresentasse limitação quanto ao número de variáveis a serem utilizadas;
- Utilizasse como base uma análise correlacional duas a duas de cada uma das variáveis, de modo a possibilitar uma análise específica das suas relações;
- Possibilitasse, de maneira prática e cientificamente comprovada, atestar a adequação do dito método aos dados obtidos para o problema específico em questão;
- Em caso de um número elevado de variáveis, possibilitasse reduzir a ordem do problema, ou seja, reduzir o número de variáveis (n) para outras componentes ou variáveis (x), com $x < n$, sem que haja perda significativa de informação dos dados ou do problema a ser solucionado;
- Fosse matematicamente e estatisticamente robusto e cientificamente testado nas mais diversas áreas do conhecimento.

Diante dos critérios estabelecidos, foi selecionado o método de análise das componentes principais (ACP) como ferramenta básica para avaliação dos pedidos de patente. Tal método permite determinar as ditas componentes principais para o problema específico de acordo com a parcela da variância geral explicada por cada uma das componentes. A partir da identificação destas novas componentes, é proposto um índice multivariado, ou Índice Geral de Complexidade (IGC) dos pedidos de patente, formado pela razão entre a soma ponderada das componentes mais significativas com seus autovalores e a soma dos próprios autovalores, autovalores estes obtidos da matriz de correlação das variáveis originais do problema. De posse destes índices de cada um dos pedidos, foi realizado um enquadramento dos mesmos em até cinco classes (Muito leve, Leve, Moderado, Pesado e Muito Pesado) de acordo com a média e o desvio padrão dos IGC obtidos. Cabe destacar que todos os passos para os cálculos estão detalhados no capítulo 4, sendo que os mesmos foram executados utilizando o *Visual Basic for Applications* (VBA), em versão do Microsoft Office 2010, linguagem de programação com ferramentas para automatização e manuseio de planilhas eletrônicas do Excel® (SENAC, 2010). Para conferência, a determinação dos autovalores e autovetores ainda foi realizada tanto manualmente quanto com o auxílio do programa Matlab e da calculadora Matrix Calculator (disponível em <https://matrixcalc.org/pt/vectors.html>).

Após a escolha do método, determinação dos índices e da forma de enquadramento dos pedidos em classes, uma última etapa desta 2ª parte é uma análise de sensibilidade/validação dos índices e enquadramento dos pedidos de patente. Para realização

desta etapa, é proposta uma pesquisa experimental/empírica visando estabelecer uma correlação entre os índices e enquadramentos encontrados e o tempo/esforço de exame dos pedidos de patente. Primeiramente, foram realizadas análises do processo substantivo de exame de patentes do INPI e do parecer padrão de exame utilizado, de modo a identificar as principais etapas de exame e variáveis direta ou indiretamente relacionadas. De posse destas principais etapas de exame, foi elaborado um formulário de pesquisa de tempo de exame para preenchimento pelos examinadores, onde são inseridas as informações de tempo de execução de cada uma das etapas. Sendo assim, é então determinada uma lista de pedidos, aqui denominada de Amostra Padrão, com tabelamento dos dados incluindo todas as variáveis de interesse além do tempo de exame. A esta nova amostra será então aplicado o método de ACP, calculados os índices IGC para cada pedido de patente, além de seus enquadramentos em classes. Neste contexto será então verificada a correlação entre os índices obtidos e o tempo de exame, além de realizada a aplicação do método de ACP incluindo diversas simulações com variações do tamanho da amostra e número de variáveis. Este procedimento visa a mostrar as variáveis que de fato impactam diretamente no tempo de exame e a representatividade do IGC perante a estas variáveis, avaliar o tamanho mínimo necessário das amostras, além de testar a aplicabilidade do método de ACP.

3.3 DISTRIBUIÇÃO DOS PEDIDOS DE PATENTE

Nesta 3ª parte da metodologia proposta foi construída uma lógica específica para distribuição dos pedidos utilizando como base, além do enquadramento nas cinco classes obtido na 2ª parte, também o enquadramento do pedido na CIP. Para tanto, primeiramente foi necessário definirmos as Zonas de Atuação Específicas (ZAE) dos examinadores, ou seja, as áreas técnicas específicas que os examinadores se sentem mais confortáveis de trabalhar de acordo com as suas experiências profissionais e formações acadêmicas. Portanto, os dados dos pedidos de patente obtidos foram separados por subclasse principal da CIP e, desta forma, foram identificadas as principais subclasses examinadas por cada um dos examinadores e foi determinada a ZAE de cada um deles. Esta zona foi obtida considerando as subclasses de pedidos de patente com ocorrências acima de 5% do total de exames realizados por cada um dos examinadores avaliados a partir da maior amostra obtida, a Amostra de Teste Inicial. É então proposta uma nova forma de distribuição dos pedidos de patente de acordo com os enquadramentos dos índices gerais e com as ZAE considerando os seguintes critérios/etapas:

- Distribuir igualmente entre os examinadores os pedidos com enquadramentos muito pesado e pesado;
- Compensar os pedidos muito pesados e pesados, respectivamente, com a distribuição de pedidos muito leves e leves;
- Distribuir os pedidos moderados restantes;
- Em todas as etapas anteriores, distribuir os pedidos de patente levando-se em conta as ZAE determinadas para cada examinador.

3.4 AVALIAÇÃO DA NOVA LÓGICA DE DISTRIBUIÇÃO

Na quarta parte da metodologia proposta, primeiramente foi obtida uma nova amostra de pedidos de patente mais recente, aqui denominada de Amostra Final de Redistribuição. Foram obtidos dados de pedidos de patente dos mesmos examinadores da área de eletricidade que compuseram a Amostra de Teste Inicial, todavia, com primeiros exames realizados entre maio e julho de 2020. Vale ressaltar que, a partir da entrada em vigor do plano de ataque ao backlog implementado pelo INPI, o processo de exame mudou para a maior parte dos pedidos de patente da área de eletricidade. Sendo assim, a obtenção desta nova amostra de redistribuição fez-se necessária para compatibilizar o processo de exame realizado pelos examinadores nela avaliados com o processo executado pelo examinador com o qual obtivemos a Amostra Padrão com tempo. Esta compatibilização foi obtida por meio de amostras padrão com tempo e de redistribuição contendo pedidos de patente do mesmo tipo, ou seja, pedidos de patente com aproveitamento de buscas de escritórios internacionais. Vale ressaltar que este tipo de pedido abrange uma média de 88% do estoque total de pedidos de patente com data de depósito até 2016 da área de eletricidade da DIRPA.

De posse desta nova amostra são aplicados o modelo e a lógica de distribuição propostos, e é então calculado, tanto para a distribuição original quanto para a nova distribuição, um Índice de Balanceamento da Distribuição (IBD), índice este que varia entre zero e um, e considera as diferenças entre as medianas das variáveis das amostras de cada examinador e as medianas gerais das variáveis da divisão. Neste contexto, destaque-se que quanto mais próximas as medianas das variáveis das amostras individuais dos examinadores estiverem das medianas gerais das variáveis da divisão, maior o valor do IBD e, provavelmente, melhor balanceada estará a distribuição. O detalhamento do IBD, incluindo explicações e análises acerca de sua formulação, limites máximo e mínimo serão apresentados no Capítulo 4 – Desenvolvimento. Em complemento à avaliação pelo IBD, serão ainda

realizados testes de hipóteses para verificar se as amostras redistribuídas de cada examinador apresentam medianas / perfis aproximadamente iguais aos da amostra total da divisão.

3.5 AVALIAÇÃO DA APLICABILIDADE EM OUTRAS ÁREAS DO INPI

Na quinta e última parte da metodologia proposta, primeiramente foram obtidas novas amostras de pedidos de patente de áreas distintas da eletricidade para verificação da aplicação do modelo desenvolvido. Como o INPI tem quatro coordenações que englobam áreas bastante específicas, foram então obtidas amostras de dados de uma divisão de cada coordenação, respectivamente das áreas de eletrônica, química (polímeros), biologia (biofármacos) e mecânica. Todas as amostras obtidas são compostas de pedidos de patente com primeiros exames realizados entre julho e dezembro de 2020 e enquadrados no plano de ataque ao backlog como tipo 6.21, ou seja, com aproveitamento de buscas realizadas por escritórios internacionais.

De posse destas novas amostras foi então aplicado o modelo desenvolvido, incluindo todos os cálculos da ACP, determinação dos índices IGC dos pedidos, enquadramento dos pedidos nas cinco classes definidas (muito leve, leve, moderado, pesado e muito pesado), aplicação da nova lógica de distribuição e cálculo dos índices IBD das distribuições original e das redistribuições.

Com a análise da aplicação do modelo para estas novas áreas e a comparação do comportamento com a aplicação já validada para a área de eletricidade, foram elaborados critérios mínimos a serem considerados, de modo a permitir uma aplicação confiável sem que seja obrigatoriamente necessário obter amostras com tempo e validações específicas para cada uma das novas áreas. Os critérios foram elaborados conforme os seguintes eixos de análise:

- Rendimento do método de ACP:
 - avaliação dos resultados para diferentes tamanhos de amostras – determinar número mínimo de amostras;
 - análise gráfica dos autovalores e variâncias acumuladas – verificar coincidência ou aproximação no comportamento dos feixes/gráficos;
 - observação dos valores de variância acumulada (ou autovalores) da componente principal Y1 – determinar valor mínimo de variância da componente Y1.

- Comportamento do índice IGC:
 - avaliação da distribuição estatística dos dados – verificar se a distribuição é aproximadamente normal;
 - avaliação da sua correlação com as variáveis de interesse – determinar correlação mínima com a principal variável de interesse.
- Funcionamento da nova lógica de distribuição:
 - avaliação do IBD antes e após a redistribuição – verificar evolução do IBD.

4 DESENVOLVIMENTO

4.1 ANÁLISE MULTIVARIADA: APLICAÇÃO DA ANÁLISE DAS COMPONENTES PRINCIPAIS (ACP)

4.1.1 Introdução

Criada por Person (1901) e depois computacionalmente descrita por Hotelling (1933, 1936), a Análise de Componentes Principais (ACP) é uma técnica estatística multivariada que tem como foco explicar a estrutura de variância-covariância de um conjunto de dados, tendo como objetivos principais a redução da dimensionalidade do problema e a melhor interpretação dos dados (JONHSON; WICHERN, 2002; RENCHER, 2002; PAIVA, 2006).

Para avaliar um pedido de patente deve-se considerar diversas variáveis medidas de cada um dos pedidos de patente da população estudada. O que o método de análise das componentes principais propõe é a aplicação de uma transformação a essas variáveis, de modo que as novas componentes obtidas possibilitem uma melhor discriminação e análise dos elementos da dita população. Segundo Herig (1979), este novo olhar é de grande valia quando se trata de criar uma tipologia para a população, executar uma classificação dos elementos segundo determinados critérios, etc.

De acordo com Rencher (2002), a ACP normalmente revela relacionamentos que não seriam previamente identificados apenas com a análise do conjunto original de dados e variáveis, possibilitando uma interpretação mais abrangente do fenômeno em estudo.

Segundo Regazzi (2001, p.1) “procura-se redistribuir a variação nas variáveis (eixos originais) de forma a obter o conjunto ortogonal de eixos não correlacionados”, sendo que o

novo conjunto de componentes mantém ao máximo a variabilidade original, ou seja, sem perda relevante de informação.

Conforme Vicini (2005, p. 29-30) “Na prática, o algoritmo baseia-se na matriz de variância-covariância, ou na matriz de correlação, de onde são extraídos os autovalores e autovetores” e “por fim, escrever as combinações lineares, que serão as novas variáveis, denominadas de componentes principais”.

Vale ressaltar que a ACP é amplamente utilizada pela comunidade científica demonstrando sua eficiência e robustez em aplicações de diversas áreas do conhecimento tais como agronomia, zootecnia, medicina, dentre outras (HONGYU, 2015; MEIRA et al, 2013; PAIVA et al, 2010; SAVENAGO et al, 2011). Exemplos de aplicações práticas da ACP para avaliação de serviços públicos de estados brasileiros, avaliação do desenvolvimento regional de municípios catarinenses e também análise de dados da criminalidade em estados americanos podem ser vistos, respectivamente, nos trabalhos de Silva, Silva e Borges (2015), Begnini e Almeida (2016) e Hongyu, Sandanielo e Junior (2015).

4.1.2 Determinação das Componentes Principais

Para a determinação das componentes principais, é necessário seguir os seguintes passos:

- ✓ Padronizar as variáveis originais de modo que todas apresentem média igual a zero e desvio padrão igual a um, evitando influência de ordens de grandeza distintas;
- ✓ Calcular a matriz de variância-covariância (S) e a matriz de correlação (R) que, no caso de variáveis padronizadas, serão iguais;
- ✓ Encontrar os autovalores das matrizes e seus respectivos autovetores;
- ✓ Selecionar as componentes calculando combinações lineares das variáveis originais com os autovetores da matriz de correlação.

Para a geração das componentes principais, deve-se montar uma matriz X de dimensão $n \times p$, onde X_1, X_2, \dots, X_p representam as variáveis originais do problema (colunas $j = 1, 2, \dots, p$) e onde n representa cada um dos indivíduos a serem observados (linhas $i = 1, 2, \dots, n$). Em nosso caso, n será o número de pedidos de patente analisados, e p o número de variáveis relevantes identificadas de cada um destes pedidos. A Tabela 1 ilustra um exemplo de tabela de dados em que cada elemento pode ser representado por X_{ij} .

Tabela 1 – Matriz X de dados de n indivíduos (pedidos de patente) com p variáveis

Indivíduos	Variáveis							
	X ₁	X ₂	X ₃	X ₄	...	X _j	...	X _p
1	X ₁₁	X ₁₂	X ₁₃	X ₁₄	...	X _{1j}	...	X _{1p}
2	X ₂₁	X ₂₂	X ₂₃	X ₂₄	...	X _{2j}	...	X _{2p}
3	X ₃₁	X ₃₂	X ₃₃	X ₃₄	...	X _{3j}	...	X _{3p}
...
...
...
i	X _{i1}	X _{i2}	X _{i3}	X _{i4}	...	X _{ij}	...	X _{ip}
...
...
n	X _{n1}	X _{n2}	X _{n3}	X _{n4}	...	X _{nj}	...	X _{np}

Fonte: Regazzi (2001) e Vicini (2005).

4.1.2.1 Padronização das Variáveis

Antes da aplicação do método, deve-se analisar os dados e verificar se as variáveis apresentam unidades e ordens de grandeza muito diferentes umas das outras e, sendo assim, executar uma padronização. Para padronização das variáveis será utilizada como referência a matriz X de dados da Tabela 1. A partir da Matriz X, é então obtida uma nova Matriz Z de dados padronizados com p variáveis (colunas j = 1, 2, ..., p) e n indivíduos (pedidos de patente em cada linha i = 1, 2, ..., n), em que cada elemento Z_{ij} é calculado segundo a função:

$$Z_{ij} = \frac{X_{ij} - \mu_p}{\delta_p} \quad (1)$$

Onde:

X_{ij} é o elemento da matriz X de dados original;

μ_p é a média da variável p, sendo p = j; e

δ_p é o desvio padrão da variável p, sendo p = j.

Sendo assim, é então obtida uma Matriz Z de dimensão n x p em que cada variável (Z₁, Z₂, ..., Z_p) agora apresenta média igual zero, além de desvio padrão e variância iguais a um.

4.1.2.2 Matrizes de Variância-Covariância (S) e de Correlação (R)

Conforme Herig (1979), variância é uma grandeza que mostra quanto uma determinada variável aleatória se afasta da sua média ou do seu valor esperado, sendo uma medida de dispersão. Já a covariância, mostra qual a tendência de variáveis Z_j e $Z_{j'}$ variarem conjuntamente, ou seja, o quanto o fato de Z_j se afastar da sua média implica em que $Z_{j'}$ também assim o faça.

Estatisticamente a variância e a covariância podem ser determinadas conforme as seguintes equações:

$$VAR[Z_j] = \frac{1}{n} \sum_{k=1}^n (Z_{jk} - \mu_j)^2 \quad (2)$$

e

$$COV[Z_j; Z_{j'}] = \frac{1}{n} \sum_{k=1}^n [(Z_{jk} - \mu_j)(Z_{j'k} - \mu_{j'})] \quad (3)$$

Onde:

$VAR[Z_j]$ é a variância da variável padronizada Z_j ;

$COV[Z_j; Z_{j'}]$ é a covariância das variáveis padronizadas Z_j e $Z_{j'}$;

n é o número de indivíduos;

μ_j e $\mu_{j'}$ são, respectivamente, as médias das variáveis padronizadas Z_j e $Z_{j'}$; e

Z_{jk} são os elementos da matriz de dados.

Já a correlação (ou coeficiente de correlação) entre duas variáveis aleatórias Z_j e $Z_{j'}$, é dada por:

$$COR[Z_j; Z_{j'}] = \frac{COV[Z_j; Z_{j'}]}{\delta_{Z_j} \cdot \delta_{Z_{j'}}} \quad (4)$$

Onde:

$COR[Z_j; Z_{j'}]$ é a correlação das variáveis padronizadas Z_j e $Z_{j'}$;

$COV[Z_j; Z_j]$ é a covariância das variáveis padronizadas Z_j e Z_j ; e

δ_j e δ_j são, respectivamente, os desvios padrão das variáveis padronizadas Z_j e Z_j .

Considerando que serão utilizadas as variáveis padronizadas (média igual a zero, desvio padrão igual a 1 e variância igual a 1) nota-se, da observação das equações (2) a (4), que a correlação passa a ser igual a covariância e, sendo assim, as matrizes de covariância (S) e de correlação (R) são iguais. Ainda observando as equações (2) e (3), verifica-se que a covariância de uma variável com ela mesma é a sua própria variância e, portanto, na diagonal principal, os elementos das matrizes de covariância deverão ser iguais a variância, ou seja, iguais a um. Além disso, como a correlação de uma variável com ela mesma é igual a um, resta confirmado que os elementos da diagonal principal serão de fato todos iguais a um. Por conseguinte, são então determinadas as matrizes S e R das variáveis padronizadas, ambas quadradas de dimensão $p \times p$, conforme as equações a seguir:

$$S = R = \begin{bmatrix} VAR[Z_1; Z_1] & COV[Z_1; Z_2] & \dots & COV[Z_1; Z_p] \\ COV[Z_2; Z_1] & VAR[Z_2; Z_2] & \dots & COV[Z_2; Z_p] \\ \vdots & \vdots & \ddots & \vdots \\ COV[Z_p; Z_1] & COV[Z_p; Z_2] & \dots & VAR[Z_p; Z_p] \end{bmatrix} \quad (5)$$

$$S = R = \begin{bmatrix} 1 & COV[Z_1; Z_2] & \dots & COV[Z_1; Z_p] \\ COV[Z_2; Z_1] & 1 & \dots & COV[Z_2; Z_p] \\ \vdots & \vdots & \ddots & \vdots \\ COV[Z_p; Z_1] & COV[Z_p; Z_2] & \dots & 1 \end{bmatrix} \quad (6)$$

4.1.2.3 Cálculo dos Autovalores e Autovetores das Matrizes S e R

De posse da matriz S, para calcularmos os autovalores ($\lambda_1, \lambda_2, \dots, \lambda_p$) de S usamos a seguinte equação proveniente da álgebra linear:

$$|S - \lambda. I| = 0 \quad (7)$$

Onde:

S é a matriz de Variância-Covariância de dimensão $p \times p$ dos dados padronizados;

λ é um dos p autovalores da matriz S;

I é a matriz identidade de ordem p ;
 p é o número total de variáveis padronizadas.

Para determinação dos autovetores (V_1, V_2, \dots, V_p) associados a cada um dos autovalores ($\lambda_1, \lambda_2, \dots, \lambda_p$) de S usamos a seguinte equação proveniente da álgebra linear:

$$S.V = \lambda.V \quad (8)$$

Onde:

S é a matriz de Variância-Covariância de dimensão $p \times p$ dos dados padronizados;

V é um autovetor de S com dimensão $p \times 1$;

λ é um dos p autovalores da matriz S ;

p é o número total de variáveis padronizadas.

4.1.2.4 Cálculo das Componentes Principais e Algumas Propriedades

Como um primeiro passo para o cálculo das componentes principais, os autovalores ($\lambda_1, \lambda_2, \dots, \lambda_p$) calculados anteriormente são colocados em ordem decrescente, de modo que $\lambda_1 > \lambda_2 > \dots > \lambda_p$. Sendo assim, é então determinada uma matriz V de dimensão $p \times p$ onde, cada coluna de 1 a p desta matriz, é um dos autovetores de S (V_1, V_2, \dots, V_p) associados, respectivamente, a cada um dos autovalores ordenados. Os elementos da matriz V são também comumente denominados de fatores ou escores das componentes principais. A equação (9) a seguir mostra a estrutura da matriz V .

$$V = [V_1 \quad V_2 \quad \dots \quad V_p] = \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1p} \\ v_{21} & v_{22} & \dots & v_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ v_{p1} & v_{p2} & \dots & v_{pp} \end{bmatrix} \quad (9)$$

Com a obtenção dos autovetores associados aos autovalores em ordem decrescente, a determinação das componentes principais (Y_1, Y_2, \dots, Y_p) para cada um dos n indivíduos em análise é determinada através da combinação linear entre as variáveis padronizadas e os autovalores calculados. Assim sendo, podemos então escrever as componentes do indivíduo n na forma da seguinte equação:

$$Y_{i(n)} = Z_{i1} \cdot v_{1i} + Z_{i2} \cdot v_{2i} + \dots + Z_{ip} \cdot v_{pi} \quad , \text{ com } i \text{ variando de } 1 \text{ até } p \quad (10)$$

Onde:

$Y_{i(n)}$ é a componente i do indivíduo n ;

Z_{ip} são os elementos da matriz Z de variáveis padronizadas; e

v_{ip} são os elementos dos autovetores calculados.

Determinadas as componentes principais (Y_1, Y_2, \dots, Y_p) dos n indivíduos de interesse, temos ainda algumas propriedades importantes:

- ✓ A variância de cada componente principal Y_i é igual ao valor do respectivo autovalor associado, ou seja: $VAR(Y_i) = \lambda_i \rightarrow VAR(Y_1) > VAR(Y_2) > \dots > VAR(Y_p)$;
- ✓ A soma das variâncias das variáveis padronizadas é igual a soma dos autovalores que é igual a soma das variâncias das componentes principais associadas, ou seja: $\sum VAR(Z_i) = \sum \lambda_i = VAR(Y_i)$;
- ✓ As covariâncias e correlações entre as componentes principais são nulas, ou seja: $COV(Y_i; Y_j) = COR(Y_i; Y_j) = 0$.

Cabe destacar que na ACP a contribuição de cada componente principal (Y_1, Y_2, \dots, Y_p) é medida em termos da sua variância. Sendo assim, podemos calcular essa contribuição por meio da relação entre a variância da componente em análise com a soma das variâncias de todas as componentes, obtendo assim a proporção (ou porcentagem) de variância total explicada por cada uma das componentes. A expressão abaixo mostra como calcular cada contribuição C_i (C_1, C_2, \dots, C_p).

$$C_i = \frac{VAR(Y_i)}{\sum_{i=1}^p VAR(Y_i)} \cdot 100 = \frac{\lambda_i}{\sum_{i=1}^p \lambda_i} \cdot 100 \quad (11)$$

Onde:

C_i é a contribuição ou % total de variância explicada pela componente Y_i ;

$VAR(Y_i)$ é a variância da componente principal Y_i ;

λ_i é um dos p autovalores da matriz S ;

p é o número total de variáveis padronizadas.

Destaque-se que maiores detalhes acerca das demonstrações matemáticas formais da ACP e suas propriedades, bem como de toda a álgebra linear envolvida, podem ser consultados nos trabalhos de Herig (1979), Johnson e Wichern (1998), Regazzi (2001) e Paiva (2006).

4.1.2.5 Critérios de Seleção das Componentes Principais

Existem diferentes critérios sendo utilizados na prática para selecionar as componentes principais de interesse. Os três principais eixos de referência para estes critérios são: os valores absolutos ou amplitudes dos autovalores associados às componentes, a representação gráfica destas amplitudes e o valor da porcentagem de variância explicada pelas componentes (vide equação 11).

O critério mais utilizado até hoje é o critério de Kaiser (1958). Nele somente são consideradas componentes principais associadas a autovalores com amplitudes maiores que a unidade, ou seja: $\lambda_i > 1$. Entretanto, alguns casos práticos mostram que apenas este critério de seleção pode não representar adequadamente a maior parte da variância total dos dados.

Uma segunda opção é realizar uma análise gráfica e verificar maiores diferenças entre os sucessivos autovalores. O critério de Cattell (1966), por exemplo, sugere que a amplitude dos autovalores seja representada graficamente em função do número dos autovalores, dispostos em ordem crescente. A seleção do número de componentes seria baseada no ponto de ruptura do gráfico. Este ponto de ruptura ocorre quando há uma queda brusca na amplitude dos autovalores (SCREMIN, 2003).

Um terceiro possível critério também bastante difundido é utilizar um valor de referência para a proporção de variância explicada pelas componentes principais. Seguindo esta lógica, as componentes principais cujo percentual acumulado de variação explicada supere o dito valor de referência devem ser selecionadas. Cabe destacar que não há consenso entre os pesquisadores sobre qual porcentagem utilizar, havendo exemplos práticos diversos. Boa parte das aplicações utiliza o valor de 70% como limite. Scremin (2003) classificou o problema em níveis de aceitação, tendo considerado razoável ou “parcialmente bom” valores compreendidos entre 62 e 80%. Algumas aplicações, como no trabalho de Johnson e Wichern (2002) utilizam valores a partir de 80%.

Embora cada critério apresente vantagens e desvantagens, nesta pesquisa preferiu-se adotar a abordagem de Souza (2000), Lopes (2001) e Paiva (2006), que utilizaram uma combinação dos três critérios descritos anteriormente.

Como valor de referência do terceiro critério, acredita-se que um percentual de variância explicada de 70%, mais especificamente entre 60 e 80%, seja um valor adequado para a seleção das componentes principais mais representativas.

4.1.3 Adequação dos Dados

4.1.3.1 Número de Variáveis e Elementos das Amostras

No que tange ao número de variáveis, não há consenso sobre um quantitativo exato a ser utilizado para realização de uma ACP. Quanto às amostras, critérios estatísticos e a experiência prática mostram que são necessárias pelo menos cerca de 50 medições ou observações totais, preferencialmente, pelo menos 100 (HAIR et al., 2009). O autor sugere ainda que haja, no mínimo, cinco observações para cada variável utilizada no estudo.

4.1.3.2 Teste de Bartlett

O teste de Bartlett permite uma comparação entre a matriz R de correlações (ou de Covariâncias) das variáveis padronizadas e a matriz identidade de igual ordem. Mais especificamente, testa se há correlações suficientes entre as variáveis de modo que possamos separá-las em fatores ou componentes. Se as variáveis forem perfeitamente correlacionadas, apenas um fator seria suficiente e, caso sejam ortogonais, isto é, com correlação zero, seriam necessários tantos fatores quanto variáveis.

Uma das variáveis relevantes para execução do teste é o determinante da matriz R. Caso ele fique próximo de um, que é o determinante da matriz identidade, tudo indica que o teste falhará. Por outro lado, quanto mais próximo de zero for o determinante de R, maiores são as correlações entre as variáveis, e maior a tendência de eficácia da aplicação da ACP. A estatística de teste é dada por:

$$\chi^2 = - \left[n - 1 - \frac{2p + 5}{6} \right] \cdot \ln|R| \quad (12)$$

Onde: $|R|$ é o determinante da matriz de correlações de dados padronizados, n é o tamanho da amostra e p o número de variáveis estudadas. A estatística de teste é aproximadamente uma qui-quadrado com $v = p \cdot (p - 1)/2$ graus de liberdade, sendo considerado um nível de significância (α) de 5%. Para aprovação no teste, o valor do qui-quadrado calculado deve ser maior que o valor teórico determinado para o nível de significância escolhido, ou seja: $\chi^2 > \chi^2_{(\alpha, v)}$.

4.1.3.3 Teste de Kaiser-Meyer-Olkin (KMO)

Um outro teste muito utilizado para verificação da adequação dos dados para realização de uma ACP é o chamado teste KMO. O teste mede a relação entre as correlações das variáveis padronizadas e as suas correlações parciais, sendo que estas últimas são uma estimativa das correlações entre as componentes ou fatores, ou seja, devem ser próximas de zero. O valor do KMO pode ser determinado pela seguinte expressão:

$$KMO = \frac{\sum_i \sum_{j \neq i} r_{ij}^2}{\sum_i \sum_{j \neq i} r_{ij}^2 + \sum_i \sum_{j \neq i} a_{ij}^2} \quad (13)$$

Onde: r_{ij}^2 é o quadrado dos elementos da matriz de correlação R fora da diagonal principal e a_{ij}^2 é o quadrado das correlações parciais entre as variáveis.

Para determinar as correlações parciais a_{ij}^2 , temos que primeiro calcular os elementos v_{ij} da inversa da matriz de correlações (R^{-1}). Sendo assim, as correlações parciais a_{ij}^2 são então obtidas pela seguinte expressão:

$$a_{ij} = - \frac{v_{ij}}{v_{ii} + v_{jj}} \quad (14)$$

Os resultados deste teste variam de 0 a 1 sendo que, quanto mais próximo de 1, melhor. Não há consenso entre autores e pesquisadores de quais valores de KMO seriam os mais adequados. Para Hair et al. (2009) são aceitáveis valores de KMO superiores a 0,5. Pallant (2007) sugere 0,6 como limite aceitável. Já Friel (2009) sugere uma escala em que 0,7 seria o limite aceitável ou mediano. No presente trabalho será adotado o limite de 0,5 como sendo o adequado.

4.1.4 Índice Geral de Complexidade (IGC)

Conforme Hongyu (2015), “De acordo com Kubrusly (2001), para se estabelecer um índice que possibilite ordenar um conjunto de n objetos, segundo critério definido por um conjunto de m variáveis adequadas, é necessário escolher os pesos ou ponderações das variáveis de tal forma que traduzam a informação contida nas variáveis”, sendo que para a construção de um índice como combinação linear de variáveis, “é desejável que este índice contenha o máximo de informação possível do conjunto de variáveis selecionadas para estudo”. Segundo Sandanielo (apud HONGYU, 2015), “um método que cria combinações lineares com máxima variância é a análise de componentes principais”.

Neste contexto, o presente trabalho utiliza um índice em uma dimensão formado a partir das componentes principais mais significativas (cuidadosamente selecionadas a partir dos critérios de seleção abordados no item 4.1.2.5), ponderadas pelos seus respectivos autovalores. Sendo assim, é então definido um Índice Geral de Complexidade (IGC) dos pedidos de patente a serem avaliados, conforme a seguinte equação:

$$IGC_n = \frac{\sum_{i=1}^k Y_i \lambda_i}{\sum_{i=1}^k \lambda_i} \quad (15)$$

Onde: Y_i são as componentes principais calculadas, λ_i são os autovalores calculados, k é o número de componentes principais selecionadas e n é o pedido de patente avaliado.

4.1.5 Enquadramento de Classes dos Pedidos de Patente

De modo a obter um agrupamento dos dados conforme o índice IGC calculado, o primeiro passo foi padronizar os valores obtidos de maneira similar ao que foi feito para os dados originais, utilizando-se a equação (1). De posse do IGC padronizado (com média igual a zero e desvio padrão igual a um) para todos os pedidos de patente em análise, a premissa considerada é de que os dados se comportam aproximadamente como uma distribuição normal, vide item 6.2.1 do portal action (disponível em <http://www.portalaction.com.br/probabilidades/62-distribuicao-normal>). O comportamento de uma distribuição normal (com média μ e desvio padrão δ) é ilustrado na figura 2 a seguir.

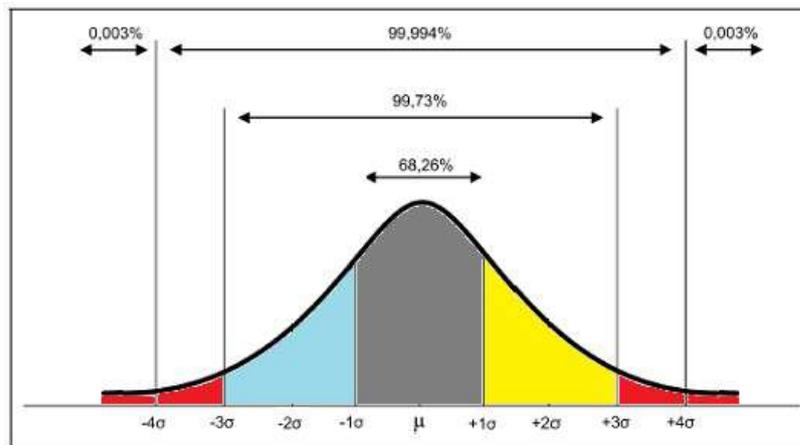


Figura 2 – Distribuição normal adaptada do portal action.

Considerando a distribuição normal da figura 2, foram então definidos como padrão três intervalos principais de largura equivalente a dois desvios padrão cada:

- Intervalo cinza: dados que se encontram relativamente perto da média, ou seja, equivalentes à média mais ou menos um desvio padrão, dentro do intervalo representado por $[\mu - \delta, \mu + \delta]$;
- Intervalo azul claro: dados que se encontram relativamente abaixo da média, ou seja, valores de média menos três desvios padrão até média menos um desvio padrão, dentro do intervalo representado por $[\mu - 3\delta, \mu - \delta]$;
- Intervalo amarelo: dados que se encontram relativamente acima da média, ou seja, valores de média mais um desvio padrão até média mais três desvios padrão, dentro do intervalo representado por $(\mu + \delta, \mu + 3\delta]$.

Foram ainda definidos mais dois intervalos relevantes conforme a figura 2:

- Intervalo vermelho à esquerda: dados que se encontram muito abaixo da média e com probabilidade muito baixa de ocorrência, ou seja, valores menores do que a média menos três desvios padrão, dentro do intervalo representado por $(-\infty, \mu - 3\delta)$;
- Intervalo vermelho à direita: dados que se encontram muito acima da média e com probabilidade muito baixa de ocorrência, ou seja, valores maiores do que a média mais três desvios padrão, dentro do intervalo representado por $(\mu + 3\delta, +\infty)$.

Cabe destacar que, para verificação da normalidade dos dados é sugerida a realização do teste estatístico de Kolmogorov-Smirnov. Mais detalhes sobre o teste podem ser vistos no item 6.2 do portal action (disponível em <http://www.portalaction.com.br/inferencia/62-teste-de-kolmogorov-smirnov>).

Sendo assim, como os dados de IGC são padronizados, ao substituir a média por zero e o desvio padrão por um, os intervalos definidos acima dão origem às faixas de enquadramento em classes conforme a Tabela 2.

Tabela 2 – Enquadramento de Classes dos Pedidos de Patente

Enquadramentos	Faixas
Muito Pesado	$IGC > 3$
Pesado	$1 < IGC \leq 3$
Moderado	$-1 \leq IGC \leq 1$
Leve	$-3 \leq IGC < -1$
Muito Leve	$IGC < -3$

Fonte: Elaboração Própria.

Após o cálculo do índice e enquadramento dos pedidos em classes, tem-se a construção do modelo completo para avaliação dos pedidos de patente. Um diagrama do modelo é apresentado na Figura 3.

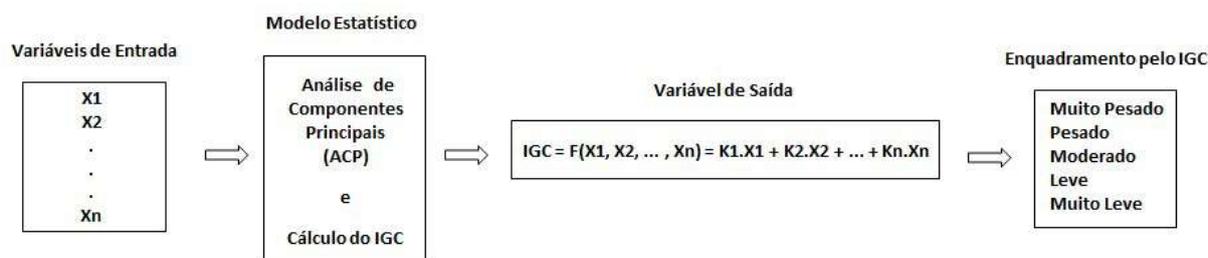


Figura 3 – Diagrama do Modelo

Sendo assim, o próximo passo é viabilizar sua avaliação por meio de uma análise de sensibilidade e das correlações com tempo de exame/análise do pedido. Esta avaliação se dará por meio da análise das etapas do exame substantivo, elaboração de um formulário de tempo de exame para obtenção de uma nova Amostra Padrão de pedidos de patente com o tempo, além da execução de diversas simulações considerando diferentes números de variáveis e tamanhos de amostra.

4.2 EXAME SUBSTANTIVO DE PATENTES

Para que o Estado conceda uma patente é necessário que a invenção atenda aos seguintes requisitos de patenteabilidade: novidade, atividade inventiva e aplicação industrial. O requisito de novidade impõe que a matéria a ser protegida não esteja compreendida no atual estado da técnica, ou seja, não tenha sido ainda revelada. O requisito de atividade inventiva exige que, para um técnico no assunto, a matéria a ser protegida não seja uma decorrência evidente ou óbvia do estado da técnica. E o requisito de aplicação industrial impõe que a invenção pleiteada apresente utilidade para a indústria e, conseqüentemente, seja passível de produção em escala industrial. No Brasil, tais requisitos do exame substantivo estão entre os pressupostos legais estabelecidos pela LPI. Cabe destacar que os pareceres dos examinadores são públicos (Via RPI – Revista de Propriedade Industrial) e o INPI-BR disponibiliza os pareceres gerados no exame técnico dos pedidos de patente, em seu sítio na internet, por meio do sistema e-Patentes/Parecer (<http://eparecer.inpi.gov.br/eparecer.php>).

4.2.1 Exame Técnico – Etapas Principais, Parecer Técnico e Variáveis Relacionadas

4.2.1.1 Etapas Principais

4.2.1.1.1 Etapa 1 – Análise Técnica Inicial

A partir do momento em que o examinador recebe um novo pedido de patente a ser examinado é evidente que a primeira etapa que requer um grande esforço de tempo é a leitura do pedido. Como o examinador precisa ler atentamente todo o pedido e acompanhar pelos desenhos, a variável relevante a ser considerada nesta etapa é o número total de páginas do pedido.

Após a leitura e entendimento do pedido, o examinador deverá realizar considerações em relação aos seguintes Artigos da LPI: Artigo 10 (não se considera invenção); Artigo 18 (não é patenteável); Artigo 22 (unidade de invenção); Artigo 32 (acréscimo de matéria); Artigo 24 (suficiência descritiva); Artigo 25 (clareza). As análises descritas fazem parte de uma primeira grande etapa do processo de exame técnico, aqui denominada de Análise Técnica Inicial.

4.2.1.1.2 Etapa 2 – Busca de Anterioridades

Após a Análise Técnica Inicial, o examinador deve realizar uma busca de anterioridades com o objetivo de determinar o estado da técnica mais próximo à matéria pleiteada. Para a realização das buscas, o examinador leva em consideração principalmente o quadro reivindicatório, mais especificamente suas reivindicações independentes. Sendo assim, o número de reivindicações independentes é uma variável relevante a ser considerada. Além disso, se estas reivindicações tiverem muitas características e forem muito extensas, haverá um esforço ainda maior. Consequentemente, o número de páginas de reivindicações também é uma variável relevante. Por fim, quanto maior o número de classificações do pedido, maior a tendência de abordar mais de uma área tecnológica ou regiões de fronteira, o que também pode tornar a busca ainda mais complexa. Portanto, o número de subclasses do pedido também é uma variável a ser considerada nesta etapa.

4.2.1.1.3 Etapa 3 – Análise Técnica Específica

Realizadas as buscas e determinadas as anterioridades úteis ao exame, a terceira grande etapa trata da comparação entre a matéria pleiteada e os conhecimentos apresentados no estado da técnica, ou seja, a análise da patenteabilidade. Esta terceira etapa é aqui denominada de Análise Técnica Específica. Quando da análise de patenteabilidade, um exame detalhado das reivindicações independentes do pedido é obrigatório. Sendo assim, a primeira variável relevante nesta etapa será o número de reivindicações independentes. Do mesmo modo que nas buscas, caso as reivindicações sejam muito extensas e apresentem muitas características técnicas, o esforço será maior. Logo, o número de páginas de reivindicações também é uma variável relevante desta etapa. Como em alguns casos as reivindicações dependentes também são analisadas em detalhe, considera-se a possibilidade do número de reivindicações dependentes ser utilizado e, neste caso, poder-se-ia usar o número total de reivindicações como variável de interesse.

4.2.1.2 Parecer Técnico do INPI – Estrutura e Preenchimento

Em seguida será realizada uma análise detalhada da estrutura do parecer técnico a ser elaborado no exame com vistas a estabelecer qual parte do mesmo pode/deve ser preenchida em cada uma das três etapas aqui definidas e quais variáveis podem ser relevantes.

Desde 2013, o INPI adota um padrão de parecer que visa auxiliar o exame substantivo. A estrutura básica do parecer está dividida em cinco quadros e uma conclusão, sendo que esta última deve explicitar os artigos da LPI que subsidiaram a decisão proferida pelo Examinador.

A Figura 4 mostra um exemplo de Quadro 1. Nele são explicitadas as páginas do processo efetivamente examinadas, apontando as petições, número de páginas de cada uma das partes do pedido, bem como a data de recebimento pelo INPI.

Quadro 1 – Páginas do pedido examinadas			
Elemento	Páginas	n.º da Petição	Data
Relatório Descritivo	1 a 27	020060078092/NPRJ	30/05/2006
Quadro Reivindicatório	1 e 2	020060078092/NPRJ	30/05/2006
Desenhos	1 a 10	020060078092/NPRJ	30/05/2006
Resumo	1	020060078092/NPRJ	30/05/2006

Comentários/Justificativas

Figura 4 – Exemplo de Quadro 1 do parecer de exame do INPI

Já o Quadro 2 prevê uma avaliação dos artigos 10, 18, 22 e 32 da LPI, de maneira que os examinadores possam analisar o escopo da invenção em questão. A Figura 5 ilustra um exemplo de Quadro 2.

Quadro 2 – Considerações referentes aos Artigos 10, 18, 22 e 32 da Lei n.º 9.279 de 14 de maio de 1996 – LPI		
Artigos da LPI	Sim	Não
A matéria enquadra-se no art. 10 da LPI (não se considera invenção)		X
A matéria enquadra-se no art. 18 da LPI (não é patenteável)		X
O pedido apresenta Unidade de Invenção (art. 22 da LPI)	X	
O pedido está de acordo com disposto no art. 32 da LPI	X	

Comentários/Justificativas

Figura 5 – Exemplo de Quadro 2 do parecer de exame do INPI

Dentre estas avaliações de escopo destaca-se a verificação se a matéria é de fato considerada uma invenção. Descobertas, concepções puramente abstratas e esquemas contábeis, por exemplo, não são considerados invenção pelo artigo 10 da LPI. Deve-se verificar também se a matéria pleiteada é patenteável ou não. Invenções consideradas contrárias a moral, ou que desejem proteger todo ou parte de seres vivos, por exemplo, não são patenteáveis de acordo com o artigo 18 da LPI. Ressalta-se que há de verificar se o pedido apresenta unidade de invenção. Ou seja, ao ler e entender o pedido, o examinador deve

identificar se o pedido apresenta apenas uma invenção. Caso contrário, deve apontar os conceitos inventivos distintos da matéria pleiteada e orientar pela necessidade de atendimento ao artigo 22 da LPI. Por fim, de acordo com o artigo 32 da LPI, não são aceitas modificações que ampliem o escopo da invenção originalmente apresentada ao INPI. Caso ocorram modificações deste gênero, principalmente no quadro reivindicatório, o Examinador deve apontar no momento do preenchimento do Quadro 2, alertando para a necessidade de atendimento ao artigo 32 da LPI.

Neste contexto, nota-se que após uma leitura atenta do pedido, verificação básica das partes do processo, além de seu entendimento técnico, já é possível o preenchimento do quadro 1 e, na maioria dos casos, também do quadro 2. Sendo assim, sugerimos o preenchimento dos quadros 1 e 2 como partes integrantes da etapa 1 do processo de exame, ou seja, da Análise Técnica Inicial. Vale destacar que, para uma análise adequada dos artigos constantes do Quadro 2, é necessária uma análise criteriosa das reivindicações do pedido. Além disso, o número de classes do pedido é um fator que indica uma tendência a ter ou não unidade de invenção. Assim sendo, as variáveis número de reivindicações independentes e número de subclasses são passíveis de serem consideradas relevantes para esta etapa.

O Quadro 3 é o local onde o examinador faz suas observações acerca dos artigos 24 e 25 da LPI, que trata das condições relacionadas ao relatório descritivo e ao quadro reivindicatório, respectivamente. No artigo 24 é abordada principalmente a suficiência descritiva do pedido, estando ela diretamente relacionada ao relatório descritivo, dado que este deve descrever a matéria pleiteada, no quadro reivindicatório, de modo que um técnico no assunto seja capaz de a realizar. Já o artigo 25 trata da clareza do quadro reivindicatório como um todo, sejam reivindicações independentes ou não. Ressalte-se que, caso haja algum problema com os desenhos, os quais auxiliam no entendimento da invenção, é no Quadro 3 que isto é apontado. Portanto, verifica-se que quanto maior for o pedido em geral e quanto mais reivindicações ele tiver, há uma tendência a um maior esforço de análise e, conseqüentemente, de um preenchimento mais detalhado do referido quadro. Sendo assim, as variáveis número total de páginas e número total de reivindicações são candidatas a serem consideradas relevantes. A Figura 6 mostra um exemplo de Quadro 3.

Quadro 3 – Considerações referentes aos Artigos 24 e 25 da LPI		
Artigos da LPI	Sim	Não
O relatório descritivo está de acordo com disposto no art. 24 da LPI		X
O quadro reivindicatório está de acordo com disposto no art. 25 da LPI	X	

Comentários/Justificativas

Figura 6 – Exemplo de Quadro 3 do parecer de exame do INPI

De maneira semelhante aos casos dos Quadros 1 e 2, o preenchimento do Quadro 3 pode ser realizado após uma leitura atenta e entendimento técnico do pedido, ou seja, pode ser enquadrado na análise técnica inicial, etapa 1 do processo de exame.

O Quadro 4, como pode ser visto na Figura 7, lista os documentos relevantes do estado da técnica que contribuem para o exame substantivo da matéria reivindicada. Obviamente seu preenchimento deriva diretamente da realização das buscas e análise de anterioridades e, sendo assim, deve ser preenchido na etapa 2.

Quadro 4 – Documentos citados no parecer		
Código	Documento	Data de publicação
D1	US6316895B1	13/11/2001
D2	US4924168A	08/05/1990

Comentários/Justificativas

Figura 7 – Exemplo de Quadro 4 do parecer de exame do INPI

Por fim, o Quadro 5 do parecer apresenta o resultado da análise dos requisitos de patenteabilidade, de acordo com os documentos de anterioridade considerados relevantes, sendo que a coluna “Reivindicações” deve contemplar todas as reivindicações analisadas. A Figura 8 mostra um exemplo de Quadro 5.

Quadro 5 – Análise dos Requisitos de Patenteabilidade (Arts. 8.º, 11, 13 e 15 da LPI)		
Requisito de Patenteabilidade	Cumprimento	Reivindicações
Aplicação Industrial	Sim	1 a 6
	Não	----
Novidade	Sim	1 a 6
	Não	----
Atividade Inventiva	Sim	1 a 6
	Não	----

Comentários/Justificativas

Figura 8 – Exemplo de Quadro 5 do parecer de exame do INPI

Cabe destacar que as análises técnicas e comparativas executadas e o consequente preenchimento detalhado das justificativas do Quadro 5 refletem, em geral, a principal parte da etapa 3, ou seja, Análise Técnica Específica.

Por fim, vale ressaltar que, ao analisar a estrutura do parecer, os procedimentos de exame e suas principais etapas, resta clara a grande influência de todas as variáveis diretamente relacionadas ao volume de dados. Além disso, nota-se particular possibilidade de influência do número de subclasses, que tenderia a impactar principalmente o tempo da fase busca de anterioridades e, conseqüentemente, leitura e análise destes documentos.

4.2.2 Formulário de Tempo de Exame

Definidas as principais etapas do exame técnico de pedidos de patente, suas particularidades em termos de preenchimento do parecer técnico no INPI, além das variáveis de interesse, foi então elaborado um formulário de tempo de exame de pedidos de patente. Os dados obtidos com seu preenchimento proporcionarão estabelecer uma correlação entre os índices e enquadramentos definidos e o tempo/esforço de exame dos pedidos de patente. As Figuras 9 e 10 mostram o modelo de formulário desenvolvido.

Formulário para Preenchimento do Tempo despendido para cada uma das Etapas do Exame Técnico												
		Etapa 1: Análise Técnica Inicial			Etapa 2: Busca de Anterioridades			Etapa 3: Análise Técnica Específica				
		Leitura e entendimento do Pedido (Preenchimento Quadros 1 a 3)			Execução da busca (Preenchimento Quadro 4)			Análise de Patentabilidade (Preenchimento Quadro 5)				
Número Pedido	Tipo	Relatório Descritivo (minutos)	Quadro Reivindicatório (minutos)	Análise Desenhos (minutos)	Total Etapa 1 (horas)	Formulação da Estratégia (minutos)	Execução da Estratégia (minutos)	Total Etapa 2 (horas)	Análise características Reivindicações (minutos)	Análise características anterioridades (minutos)	Comparação e Justificativas Técnicas (minutos)	Total Etapa 3 (horas)
1					0			0				0
2					0			0				0
3					0			0				0
4					0			0				0
5					0			0				0
6					0			0				0
7					0			0				0
8					0			0				0
9					0			0				0
10					0			0				0

Figura 9 – Formulário de Tempo de Exame (Parte 1)

Avaliação do grau de esforço/dificuldade			
De 0 a 10, preencher que nota daria para o pedido (Quanto mais próximo de 10 mais difícil)			
Número Pedido	Tempo Total Pedido (horas)	Nota (0 a 10)	Espaço para Justificativa (caso ache necessário)
1	0		
2	0		
3	0		
4	0		
5	0		
6	0		
7	0		
8	0		
9	0		
10	0		

Figura 10 – Formulário de Tempo de Exame (Parte 2)

Cabe destacar que, além de considerar e inserir o campo “Tempo Total Pedido” para obtenção do tempo total de exame, nossa principal variável de interesse, foram ainda inseridos campos para identificação do tempo por etapas de exame 1 a 3, além de campos parciais dentro de cada uma das etapas, de modo a possibilitar análises específicas posteriores.

Vale ressaltar ainda que atualmente está em vigor o plano de ataque ao *backlog* anunciado pelo INPI em 2019 (INPI, 2019) e com previsão de duração até 2021. No âmbito deste plano, há pedidos específicos em que são executados exames preliminares (6.20 e 6.21), onde em muitos casos pode haver o aproveitamento de buscas já executadas em escritórios internacionais. Além disso, quanto aos pedidos sem buscas de outros escritórios, em caso de leitura e busca com identificação de anterioridades relevantes, deve ser emitido um parecer preliminar (6.22) sem a realização da etapa 3 (Análise Técnica Específica). Sendo assim, também foi incluído um campo “Tipo” para identificação destes tipos específicos de pedidos de patente, de modo que possam ser avaliados os seus efeitos com relação ao tempo de exame.

4.3 NOVA METODOLOGIA PARA DISTRIBUIÇÃO DOS PEDIDOS DE PATENTE

4.3.1 Determinação das Zonas de Atuação Específicas (ZAE) dos Examinadores

Os dados de todos os pedidos de patente em análise foram organizados separadamente por subclasse principal da CIP e por examinador. Desta forma foram identificadas as principais subclasses examinadas por cada um dos examinadores e foi determinada a Zona de Atuação Específica (ZAE) de cada um deles. Esta zona de atuação foi obtida considerando as subclasses de pedidos de patente em análise com ocorrências acima de 5% do total de exames realizados por cada um dos examinadores avaliados.

4.3.2 Lógica Iterativa para Distribuição Individual dos Pedidos

Após o enquadramento de todos os pedidos a serem distribuídos no período selecionado, é então proposta uma sequência iterativa de passos para sua distribuição buscando priorizar a escolha de pedidos respeitando, sempre que possível, além das ZAE de cada um dos examinadores, o enquadramento pelo IGC de cada um dos pedidos em distribuição. A sequência de passos a ser realizada é descrita a seguir:

1. Determinar o número de pedidos e quais os pedidos que estão enquadrados em cada uma das cinco classes definidas (muito leve, leve, moderado, pesado e muito pesado);
2. Distribuir os pedidos muito pesados:
 - 2.1. Ordenar os pedidos muito pesados por valores decrescentes de IGC;
 - 2.2. Selecionar o pedido muito pesado atual a ser distribuído;
 - 2.3. Verificar a subclasse principal da CIP do pedido muito pesado atual:
 - 2.3.1. Se a CIP do pedido atual não pertencer à ZAE de nenhum examinador, distribuir o pedido para o examinador com o menor número de pedidos muito pesados e ir para o próximo pedido (passo 2.2);
 - 2.3.2. Se a CIP do pedido atual pertencer à ZAE de um ou mais examinadores, distribuir o pedido para o examinador com o menor número de pedidos muito pesados e ir para o próximo pedido (passo 2.2);
 - 2.4. Repetir os passos 2.2 e 2.3 até que todos os pedidos muito pesados tenham sido distribuídos e, assim sendo, ir para o passo 3;
3. Distribuir os pedidos muito leves compensando, na medida do possível, os pedidos muito pesados já distribuídos no passo 2:
 - 3.1. Selecionar o pedido muito leve atual a ser distribuído;
 - 3.2. Verificar a subclasse principal da CIP do pedido muito leve atual;
 - 3.3. Determinar a diferença entre o número de pedidos muito pesados e muito leves distribuídos a cada um dos examinadores:
 - 3.3.1. Se houver diferenças positivas (indicação de necessidade de compensar), verificar se o pedido muito leve atual pertence à ZAE dos examinadores a compensar, e distribuir para o de maior diferença;
 - 3.3.1.1. Se o pedido não pertencer à ZAE de nenhum examinador, distribuir para o de maior diferença;
 - 3.3.2. Se não houver diferença positiva (indicação de que não há necessidade de compensar), distribuir o pedido para o examinador com menor número de pedidos já distribuídos até o momento;
 - 3.4. Repetir os passos 3.1 a 3.3 até que todos os pedidos muito leves tenham sido distribuídos;
4. Distribuir todos os pedidos pesados de maneira análoga à distribuição dos pedidos muito pesados (lógica similar ao passo 2);

5. Distribuir todos os pedidos leves, compensando os pedidos pesados distribuídos no passo 4, de maneira análoga à distribuição dos pedidos muito leves (lógica similar ao passo 3);
6. Distribuir todos os pedidos moderados restantes de maneira análoga aos muito pesados (passo 2) e pesados (passo 4), porém, considerando a restrição de que o número total de pedidos distribuídos a cada um dos examinadores deve ser o mesmo.

4.3.3 Amostra Final de Redistribuição e Cálculo do Índice de Balanceamento da Distribuição (IBD)

Para realização da avaliação da distribuição dos pedidos de patente foi obtida uma nova amostra de dados, denominada Amostra Final de Redistribuição. Como a Amostra Padrão (com o tempo) utilizada para validação do modelo foi obtida já com o plano de combate ao *backlog* em vigor, todos os pedidos de patente desta amostra obtidos foram do tipo 6.21, ou seja, com aproveitamento de buscas de escritórios internacionais. Sendo assim, a obtenção desta nova amostra de redistribuição fez-se necessária para compatibilizar o processo de exame realizado pelos examinadores nela avaliados com o processo executado pelo examinador com o qual obteve-se a Amostra Padrão com tempo.

De posse desta nova amostra são aplicados o modelo e a lógica de distribuição propostos, e é então calculado, tanto para a distribuição original quanto para a nova distribuição, um Índice de Balanceamento da Distribuição (IBD), conforme a equação 16:

$$IBD = \frac{\sqrt{\sum_i \sum_j \left(\frac{m_{ed_{ij}}}{M_{ed_j}}\right)^2}}{\sqrt{\sum_i \sum_j \left(\frac{m_{ed_{ij}}}{M_{ed_j}}\right)^2 + \sum_i \sum_j \left(\frac{m_{ed_{ij}} - M_{ed_j}}{M_{ed_j}}\right)^2}} \quad (16)$$

Onde:

$m_{ed_{ij}}$ é a mediana da variável j dos pedidos de patente distribuídos ao examinador i ; e
 M_{ed_j} é a mediana da variável j de todos os pedidos da amostra final de redistribuição.

Em uma primeira análise da equação do IBD, verifica-se que o índice procura captar e medir a influência das diferenças entre as medianas das variáveis das amostras de cada examinador e as medianas gerais das variáveis da divisão (amostra completa). Vale ressaltar que todas as medianas das variáveis que compõem o IBD são normalizadas (divididas) pelos valores gerais das respectivas medianas da amostra completa. Com isso, busca-se evitar maiores distorções devido a diferentes ordens de grandeza de determinadas variáveis.

Adicionalmente, como as diversas medianas calculadas podem ser maiores ou menores do que a respectiva mediana da divisão, as diferenças destes valores podem ser positivas ou negativas. Sendo assim, como deseja-se obter uma medição acumulada de todas as diferenças de medianas sem perda de informação e sem que um desvio negativo em determinada variável compense um desvio positivo de outra, optou-se por elevar ao quadrado as diferenças e, em seguida, somar e extrair a raiz quadrada.

Mais especificamente, quando são analisados o numerador e o denominador da equação do IBD, nota-se que ambos possuem um primeiro termo idêntico, que é a soma dos quadrados das medianas normalizadas de todas as variáveis de interesse das amostras dos examinadores. Todavia, o denominador apresenta um segundo termo adicional apresentando a soma dos quadrados das diferenças normalizadas entre as medianas das amostras individuais de cada examinador e as medianas das respectivas variáveis da amostra completa.

Cabe destacar que, em uma distribuição ideal, as medianas das variáveis das amostras de todos os examinadores seriam iguais as medianas das variáveis gerais da divisão, ou seja, a soma dos quadrados das diferenças das medianas (segundo termo do denominador do IBD) seria zero e, conseqüentemente, o IBD seria igual a 1. Por outro lado, uma distribuição aleatória em que as amostras dos examinadores apresentem grandes diferenças nos valores de medianas quando comparadas às medianas gerais da divisão, levaria a valores muito altos do denominador, fazendo com que o IBD fosse bastante reduzido e, no limite, tendesse a zero. Por conseguinte, temos que, quanto mais próximas as medianas das variáveis das amostras individuais dos examinadores estiverem das medianas gerais das variáveis da divisão, maior e mais próximo de 1 estará o valor do IBD e, conseqüentemente, melhor balanceada estará a distribuição.

5 RESULTADOS

Neste capítulo serão apresentados os resultados obtidos com a pesquisa. Neste contexto vale ressaltar que, para execução da metodologia proposta, em um primeiro momento foi necessária a obtenção de três amostras de dados referentes à área de eletricidade, as quais são complementares, mas não são necessariamente comparáveis entre si, pois inclusive apresentam diferentes objetivos específicos. A amostra de teste inicial, uma amostra bem grande e anterior ao plano de ataque ao *backlog*, será utilizada para uma avaliação mais geral do perfil de dados, viabilização da identificação das ZAE dos examinadores e também para uma avaliação inicial da adequação ou não dos dados ao método de análise multivariada escolhido, a ACP. A execução de teste de redistribuição com a amostra inicial se dará por questões didáticas e para verificar a aplicabilidade do método passo a passo, possibilitando uma maior discussão. A segunda amostra da área de eletricidade analisada, a amostra padrão com medição de tempo, foi obtida com o plano de ataque ao *backlog* em vigor, sendo utilizada para testes de sensibilidade da aplicação da ACP e do modelo proposto, viabilizando sua validação por meio de uma série de simulações com análise das correlações dos índices e variáveis com o tempo. Foi ainda obtida uma terceira amostra de dados da área de eletricidade, a amostra final de redistribuição, está com um perfil de pedidos similares ao da amostra padrão, possibilitando uma aplicação do modelo proposto e execução de testes reais de redistribuição sem a necessidade de medição de tempo.

Adicionalmente, foram obtidas quatro novas amostras de pedidos de patente de áreas distintas da eletricidade para uma análise mais ampla da aplicação do modelo desenvolvido. Foram então obtidas amostras de dados de uma divisão de cada coordenação do INPI-BR, respectivamente das áreas de eletrônica, química (polímeros), biologia (biofármacos) e mecânica. Por meio da análise da aplicação do modelo para estas novas áreas e da comparação do comportamento com a aplicação já validada para a área de eletricidade, o objetivo foi o de conseguir elaborar critérios mínimos a serem considerados, de modo a permitir uma aplicação confiável do modelo sem que seja obrigatoriamente necessário obter amostras com tempo e validações específicas para cada uma das novas áreas.

Sendo assim, os resultados serão apresentados em quatro partes principais referentes respectivamente a: amostra de teste inicial, amostra padrão com tempo, amostra final de redistribuição e amostras das demais áreas (distintas da eletricidade).

5.1 AMOSTRA DE TESTE INICIAL

5.1.1 Variáveis de Interesse Analisadas

Conforme item 2.6, foram identificadas 12 (doze) possíveis variáveis a serem utilizadas no presente estudo. Todavia, boa parte dos pedidos de patente depositados no INPI só incluem as citações de estado da técnica após um primeiro exame já realizado, ou seja, após o pedido já ter sido distribuído, sendo que citações de outros documentos ainda exigiriam acesso a mais de uma base de dados. Neste contexto, considera-se que estas variáveis não atendem plenamente aos critérios de “momento de disponibilidade” e “confiabilidade e eficiência”, ambos descritos no capítulo 3 – item 3.1. Sendo assim, as variáveis referentes a citações serão desconsideradas, sendo então utilizadas 10 (dez) variáveis para obtenção da amostra de teste inicial, quais sejam:

- ✓ número de páginas do relatório descritivo (variável 1);
- ✓ número de páginas de reivindicações (variável 2);
- ✓ número de páginas de desenhos (variável 3);
- ✓ número de páginas de subsídios (variável 4);
- ✓ número de reivindicações independentes (variável 5);
- ✓ número de reivindicações dependentes (variável 6);
- ✓ número de subclasses da CIP (variável 7);
- ✓ ano de depósito (variável 8);
- ✓ número de inventores (variável 9); e
- ✓ número de prioridades (variável 10).

Como a amostra de teste inicial não apresenta medição de tempo do exame, de modo a identificar as variáveis diretamente relacionadas ao volume de dados do pedido de patente, foram calculadas as correlações de todas as dez variáveis acima com o número total de páginas e de reivindicações. Os resultados evidenciam concordância com a hipótese inicial 1, tendo sido, em princípio, identificadas as seguintes variáveis de volume direto:

- ✓ número de páginas do relatório descritivo (correlação 0,88 com total de páginas);
- ✓ número de páginas de reivindicações (correlação 0,82 com total de reivindicações);
- ✓ número de páginas de desenhos (correlação 0,71 com total de páginas);

- ✓ número de reivindicações independentes (correlação 0,61 com total de reivindicações);
- ✓ número de reivindicações dependentes (correlação 0,99 com total de reivindicações);

As demais cinco variáveis tiveram correlações de volume positivas, porém, bem baixas, sugerindo que elas podem ter uma contribuição de volume de dados residual.

5.1.2 Análise da Amostra de Teste Inicial

Para esta análise foi selecionada a área tecnológica referente à engenharia elétrica, mais especificamente a Divisão de Eletricidade da DIRPA. Foram seguidos estritamente os passos da metodologia proposta no que concerne à seleção do marco temporal e demais premissas já definidas quanto ao arcabouço documental a ser analisado.

Verificou-se que a Divisão de Eletricidade possui um total de dezenove (19) Pesquisadores em Propriedade Industrial sendo que, excluindo os cargos de chefia, além de casos específicos de Pesquisadores de licença e em treinamento/tutoria ou estágio probatório, foram obtidos dados de um total de onze (11) Examinadores a serem plenamente analisados.

Para cada um dos pedidos de patente dos referidos examinadores com pelo menos um exame já realizado foram obtidas todas as variáveis de interesse, perfazendo um total de oitocentos e quatorze (814) pedidos de patente a serem avaliados e compondo a amostra de teste inicial. A Tabela 3 mostra a estrutura dos dados da amostra de teste inicial.

Tabela 3 – Estrutura da Amostra de Teste Inicial

Item	Examinador	Pedido	Subclasse Principal	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Páginas de Subsídios	Reivindicações Independentes	Reivindicações Dependentes	Subclasses	Ano Depósito	Inventores	Prioridades
1	1	PI0101329-7	H01C	24	3	5	0	1	13	1	1	7	1
2	1	PI0111250-3	H01H	7	3	2	0	1	7	1	1	2	1
3	1	PI0400832-4	H02K	12	2	2	0	3	3	1	4	1	0
4	1	PI0403974-2	H01B	40	11	21	0	9	44	2	4	5	0
5	1	PI0107166-1	H01H	16	2	1	0	2	10	1	1	4	1
6	1	PI0303675-8	H02G	7	2	4	0	1	12	1	3	3	1
.
.
.
814	11	PI0601239-6	H05B	18	2	10	0	4	3	1	6	2	0

Fonte: Elaboração Própria.

Com relação à tabela 3, cabe destacar que os pedidos da amostra apresentam anos de depósito desde 2000 até 2016, sendo que para maior conveniência de cálculos e apresentação dos dados, consideramos tais anos de depósito como, respectivamente, 0 até 16, conforme ilustrado na coluna “Ano Depósito”.

A Tabela 4 mostra, para o conjunto total de 814 pedidos, um quadro geral das principais variáveis obtidas e os resultados dos cálculos de estatísticas básicas tais como: média, desvio padrão, mediana, moda, mínimo, máximo e coeficiente de variação.

Tabela 4 – Estatísticas básicas da Amostra de Teste Inicial

	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Páginas de Subsídios	Reivindicações Independentes	Reivindicações Dependentes	Subclasses	Ano Depósito	Inventores	Prioridades
Média	16,45	4,19	5,99	0,50	2,25	13,83	1,41	4,32	2,40	0,83
Desvio Padrão	14,35	3,31	6,08	8,95	1,70	12,19	0,68	2,49	1,80	0,67
Mediana	12	3	4	0	2	11	1	4	2	1
Moda	8	2	2	0	1	8	1	4	1	1
Mínimo	2	1	0	0	1	0	1	0	1	0
Máximo	182	26	97	233	11	88	6	16	29	7
Coefficiente de Variação	87,25%	78,95%	101,61%	1781,22%	75,33%	88,15%	48,54%	57,74%	74,91%	80,25%

Fonte: Elaboração Própria.

Da análise da Tabela 4 se verifica que, seja qual for a variável do pedido a ser considerada, os desvios padrão (e variância, pois é o quadrado do desvio) e coeficientes de variação (relação entre o desvio padrão e a média) são extremamente altos, partindo de cerca de 50% para o número de classes do pedido, até perto de 100% ou mais para outras variáveis. Além disso, modas, médias e medianas são distintas entre si, sendo que os valores mínimos e máximos são bastante diferentes. Isso evidencia que duas características intrínsecas dos pedidos analisados são as suas grandes amplitude e dispersão, ou seja, há uma variabilidade grande entre eles. Tal fato mostra que, caso a distribuição destes pedidos por examinador não esteja adequada, há possibilidade de distorções e desbalanceamentos na carga dos examinadores. Neste contexto, já se verifica que uma análise que leve em conta essa grande dispersão se faz necessária, sendo que uma avaliação em termos das variâncias parece promissora.

Destaque-se que, como os dados das variáveis apresentam alta amplitude, com diferenças bem grandes entre valores máximos e mínimos, além de altas variâncias, as médias não necessariamente são boas estatísticas para uma representação e avaliação dos dados em maior detalhe. Ressalte-se ainda que, para a maioria das variáveis, as modas ficaram mais próximas das medianas do que das médias. Portanto, uma avaliação considerando as medianas dos dados evidencia-se como sendo adequada.

A Tabela 5 mostra as medianas da amostra de teste inicial com os dados separados por ano de depósito do pedido. Nesta tabela foram hachurados: na cor cinza claro as variáveis dos anos com mediana abaixo da mediana geral dos dados; e na cor cinza escuro as variáveis dos anos com mediana acima da mediana geral dos dados. Foi dado um destaque particular aos

anos de 2001 a 2006, pois foram os anos base efetivamente examinados nos anos em análise, os demais sendo casos esporádicos, com amostras bem pequenas.

Tabela 5 – Medianas da Amostra de Teste Inicial separadas por ano de depósito do pedido

	Ano Depósito	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Páginas de Subsídios	Reivindicações Independentes	Reivindicações Dependentes	Nr. de Subclasses	Nr. de Inventores	Nr. de Prioridades	Nr. de Pedidos
Médianas	geral	12,00	3,00	4,00	0,00	2,00	11,00	1,00	2,00	1,00	814
	2000	6,00	3,00	4,00	0,00	2,00	8,00	1,00	2,00	1,00	5
	2001	12,00	3,00	4,00	0,00	1,00	9,00	1,00	2,00	1,00	52
	2002	13,50	4,00	5,00	0,00	2,00	11,50	1,00	2,00	1,00	86
	2003	13,00	3,50	4,00	0,00	2,00	11,50	1,00	2,00	1,00	158
	2004	13,00	3,00	5,00	0,00	2,00	12,00	1,00	2,00	1,00	189
	2005	12,00	3,00	4,00	0,00	2,00	10,00	1,00	2,00	1,00	181
	2006	11,00	3,00	4,00	0,00	2,00	11,00	1,00	2,00	1,00	89
	2007	24,50	7,50	8,00	0,00	2,00	16,00	2,50	2,50	1,00	8
	2008	13,50	4,00	4,00	0,00	2,00	13,50	1,50	2,00	0,00	6
	2009	4,00	1,00	2,00	0,00	1,00	1,00	1,00	1,00	0,00	3
	2010	15,00	4,00	5,00	0,00	1,00	8,00	1,00	2,00	0,00	3
	2011	8,00	3,00	2,00	0,00	2,00	9,00	2,00	2,00	0,00	3
	2012	19,00	2,00	2,00	0,00	2,00	4,00	1,00	4,00	1,00	3
	2013	7,50	2,00	4,00	0,00	1,00	2,00	2,00	2,00	0,00	10
	2014	16,00	2,50	3,50	0,00	2,00	8,00	1,00	1,00	0,00	8
	2015	9,00	2,50	6,00	0,00	1,00	7,00	2,00	3,00	0,00	8
2016	54,50	3,00	54,50	0,00	1,50	7,50	1,50	16,00	0,00	2	

Fonte: Elaboração Própria.

Da análise da Tabela 5, particularmente dos anos de 2001 a 2006, verifica-se que os pedidos dos anos de 2002 a 2004 foram os de maior mediana, tendendo a serem os de maior volume de dados a serem analisados. Ainda assim, as medianas não tiveram tanta diferença, mostrando que em termos de volume, embora haja alguma influência da variável ano de depósito, esta não parece ser significativa. Nota-se ainda que, ao contrário do senso comum muito preconizado, os pedidos não parecem ter uma tendência a serem necessariamente maiores quanto mais recentes os seus anos de depósito.

A Tabela 6 mostra as medianas da amostra de teste inicial com os dados separados por examinador para quem o pedido foi distribuído. Nesta tabela foram hachurados: na cor cinza claro as variáveis dos pedidos de examinadores com mediana abaixo da mediana geral dos dados; e na cor cinza escuro as variáveis dos pedidos dos examinadores com mediana acima da mediana geral dos dados.

Tabela 6 – Medianas da Amostra de Teste Inicial separadas por examinador

	Examinador	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Páginas de Subsídios	Reivindicações Independentes	Reivindicações Dependentes	Subclasses	Ano Depósito	Inventores	Prioridades
Medianas	1	12,00	3,00	5,00	0,00	2,00	12,50	1,00	4,00	2,00	1,00
	2	11,50	3,00	4,00	0,00	1,50	11,50	1,00	4,00	2,00	1,00
	3	12,00	3,00	4,00	0,00	1,00	9,50	1,00	4,00	2,00	1,00
	4	14,50	3,00	6,00	0,00	1,50	11,00	1,00	4,00	3,00	1,00
	5	15,00	4,00	5,00	0,00	2,00	12,00	1,00	4,00	2,00	1,00
	6	13,00	4,00	4,00	0,00	2,00	12,00	1,00	4,00	2,00	1,00
	7	10,00	3,00	4,00	0,00	2,00	11,00	1,00	3,00	2,00	1,00
	8	14,00	3,00	4,00	0,00	2,00	10,50	1,00	3,00	2,00	1,00
	9	12,00	3,00	4,00	0,00	2,00	9,00	1,00	5,00	2,00	1,00
	10	12,00	3,00	5,00	0,00	1,00	10,00	1,00	4,00	2,00	1,00
	11	11,00	3,00	4,00	0,00	2,00	9,00	1,00	4,00	2,00	1,00
	Geral	12,00	3,00	4,00	0,00	2,00	11,00	1,00	4,00	2,00	1,00

Fonte: Elaboração Própria.

Da análise da Tabela 6, verifica-se que há bastantes variáveis de pedidos dos examinadores com mediana tanto abaixo quanto acima da mediana geral dos dados da divisão. Todavia, nota-se que os examinadores de números 5 e 6 parecem ter uma tendência a estarem sobrecarregados em relação aos demais, enquanto os examinadores de números 7 e 11 parecem estar com carga abaixo dos demais.

5.1.3 Execução da Análise de Componentes Principais

5.1.3.1 Padronização da Amostra de Teste Inicial

Os dados da Amostra de Teste Inicial (vide estrutura da Tabela 3) foram padronizados conforme a equação (1). A Tabela 7 mostra a estrutura dos dados da amostra de teste inicial com dados padronizados.

Tabela 7 – Estrutura da Amostra de Teste Inicial com dados padronizados

Item	Examinador	Pedido	Subclasse Principal	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Páginas de Subsídios	Reivindicações Independentes	Reivindicações Dependentes	Subclasses	Ano Depósito	Inventores	Prioridades	IGC
1	4	PI0005571-9	H02K	-0,80	-0,56	-0,33	-0,06	0,44	-0,72	-0,60	-1,73	2,56	0,25	-0,10
2	7	PI0015910-7	H01R	0,53	0,85	0,82	-0,06	1,03	1,65	-0,60	-1,73	-0,22	0,25	1,29
3	8	PI0003903-9	H02P	0,32	1,15	0,33	-0,06	-0,74	-0,15	-0,60	-1,73	-0,22	0,25	0,35
4	8	PI0015017-7	F02D	-0,73	-0,36	-0,66	-0,06	-0,15	-0,48	-0,60	-1,73	-0,78	0,25	-0,55
5	9	PI0004461-0	H05K	-0,73	-0,96	-0,82	-0,06	-0,74	-1,13	-0,60	-1,73	-0,78	-1,25	-1,26
6	1	PI0101329-7	H01C	0,53	-0,36	-0,16	-0,06	-0,74	-0,07	-0,60	-1,33	2,56	0,25	0,22
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
814	9	BR122014006372-5	B60C	5,89	1,15	-0,82	-0,06	1,03	3,13	2,32	3,88	3,68	-1,25	2,81
			Média	0	0	0	0	0	0	0	0	0	0	0
			Desvio Padrão	1	1	1	1	1	1	1	1	1	1	1
			Variância	1	1	1	1	1	1	1	1	1	1	1

Fonte: Elaboração Própria.

5.1.3.2 Cálculo da Matriz de Correlação e Seleção das Componentes Principais

Após a padronização dos dados, com a aplicação das equações (5) ou (6), são então determinadas as matrizes S e R das variáveis padronizadas e, como são dez as variáveis utilizadas, ambas as matrizes são iguais e quadradas de dimensão 10 x 10, conforme mostrado a seguir.

$$S = \begin{bmatrix} 1 & 0,47 & 0,61 & -0,03 & 0,27 & 0,35 & 0,09 & 0 & 0,28 & 0,19 \\ 0,47 & 1 & 0,28 & -0,05 & 0,58 & 0,80 & 0,04 & -0,12 & 0,13 & 0,22 \\ 0,61 & 0,28 & 1 & -0,03 & 0,17 & 0,20 & 0 & 0,07 & 0,39 & 0,08 \\ -0,03 & -0,05 & -0,03 & 1 & -0,04 & -0,05 & -0,03 & 0,06 & -0,04 & -0,07 \\ 0,27 & 0,58 & 0,17 & -0,04 & 1 & 0,51 & 0,03 & -0,08 & 0,11 & 0,20 \\ 0,35 & 0,80 & 0,20 & -0,05 & 0,51 & 1 & -0,02 & -0,13 & 0,12 & 0,20 \\ 0,09 & 0,04 & 0 & -0,03 & 0,03 & -0,02 & 1 & 0,11 & 0,06 & -0,01 \\ 0 & -0,12 & 0,07 & 0,06 & -0,08 & -0,13 & 0,11 & 1 & 0,05 & -0,25 \\ 0,28 & 0,13 & 0,39 & -0,04 & 0,11 & 0,12 & 0,06 & 0,05 & 1 & 0,07 \\ 0,19 & 0,22 & 0,08 & -0,07 & 0,20 & 0,20 & -0,01 & -0,25 & 0,07 & 1 \end{bmatrix}$$

Tendo obtido as matrizes S e R foram então aplicadas as equações (7) a (9) para determinação dos autovalores ($\lambda_1, \lambda_2, \dots, \lambda_{10}$) e da matriz V de autovetores associados (escores das componentes), os quais são mostrados a seguir.

$$\lambda = (2,95 \ 1,48 \ 1,12 \ 1,03 \ 0,92 \ 0,73 \ 0,71 \ 0,52 \ 0,34 \ 0,18)$$

e

$$V = \begin{bmatrix} 0,41 & -0,32 & -0,11 & -0,05 & 0,08 & 0,44 & -0,15 & -0,01 & 0,68 & 0,15 \\ 0,50 & 0,19 & 0,24 & -0,01 & -0,07 & 0,01 & -0,07 & 0,26 & -0,01 & -0,76 \\ 0,33 & -0,49 & -0,22 & -0,16 & -0,05 & 0,26 & -0,13 & -0,21 & -0,67 & -0,01 \\ -0,06 & -0,06 & 0,34 & -0,67 & 0,65 & -0,08 & -0,02 & 0,01 & -0,02 & 0,00 \\ 0,39 & 0,22 & 0,26 & 0,04 & -0,05 & -0,22 & 0,18 & -0,79 & 0,06 & 0,11 \\ 0,46 & 0,25 & 0,25 & -0,04 & -0,13 & -0,09 & -0,05 & 0,47 & -0,20 & 0,62 \\ 0,03 & -0,20 & 0,30 & 0,71 & 0,55 & -0,02 & -0,22 & 0,00 & -0,12 & 0,03 \\ -0,09 & -0,46 & 0,50 & 0,04 & -0,23 & 0,15 & 0,67 & 0,13 & 0,02 & -0,01 \\ 0,22 & -0,44 & -0,27 & 0,00 & 0,02 & -0,79 & 0,08 & 0,14 & 0,15 & -0,03 \\ 0,22 & 0,26 & -0,47 & 0,11 & 0,44 & 0,15 & 0,65 & 0,10 & -0,09 & -0,01 \end{bmatrix}$$

De posse dos autovalores e autovetores, são então aplicadas as equações (10) e (11) para determinação das componentes principais e das suas respectivas contribuições de

variâncias. É ainda calculada a contribuição acumulada das componentes (% de variância acumulada). A Tabela 8 mostra estes resultados.

Tabela 8 – Autovalores e Contribuições de Variância das Componentes Principais

Componente Principal	Autovalores	Variância	Variância Acumulada
Y1	2,95	29,53%	29,53%
Y2	1,48	14,82%	44,36%
Y3	1,12	11,18%	55,53%
Y4	1,03	10,29%	65,82%
Y5	0,92	9,23%	75,05%
Y6	0,73	7,33%	82,39%
Y7	0,71	7,12%	89,51%
Y8	0,53	5,25%	94,77%
Y9	0,34	3,43%	98,19%
Y10	0,18	1,81%	100,00%

Fonte: Elaboração Própria.

Da análise da Tabela 8, no caso de utilização apenas do critério considerando autovalores maiores do que um, seriam selecionadas apenas as 4 (quatro) primeiras componentes. Adicionalmente, verifica-se que estas seriam responsáveis por 65,82% da variância total. Verifica-se ainda que há uma queda mais brusca de amplitude ao chegar ao autovalor da componente 6 (0,73). Além disso, as primeiras cinco componentes explicam uma variância total de 75,05%. Sendo assim, foram então selecionadas estas cinco primeiras componentes para realização dos próximos passos.

5.1.3.3 Testes de Adequação dos Dados

Antes de proceder ao cálculo completo e a uma análise mais detalhada das componentes principais, foram realizados os testes de Bartlett e KMO. Para os cálculos foram aplicadas as equações (12) a (14).

5.1.3.3.1 Teste de Bartlett

O determinante da matriz de correlações foi calculado, sendo $\text{Det}(R) = 0,076$. Como pode ser visto um valor bem próximo de zero, tal como desejado.

O número de amostras é $n = 814$ e o número de variáveis $p = 10$. Utilizando a equação (12) foi então calculada a estatística de teste, onde: $\chi^2 = 2085,518$.

Para determinar os graus de liberdade do qui-quadrado teórico usamos a expressão $\nu = p \cdot (p - 1)/2$, onde teremos então que $\nu = 45$. Considerando um nível de confiança de 5%, temos então que $\chi^2_{(\alpha,\nu)} = \chi^2_{(0,05,45)} = 61,65$. Portanto, temos que $\chi^2 \gg \chi^2_{(\alpha,\nu)}$, ou seja, obtemos aprovação no teste.

5.1.3.3.2 Teste KMO

Conforme a equação (13), a partir da matriz de correlações R foram somados os quadrados de todos os elementos fora da diagonal principal de modo a obter o numerador da expressão do KMO. O valor encontrado foi de 5,596. Para obtenção do valor do denominador é necessário determinar a inversa da matriz R para obter as correlações parciais conforme a equação (14). Foram somados os quadrados de todas as correlações parciais com o valor já obtido para o numerador. Deste modo foi encontrado um valor de 7,965 para o denominador. Sendo assim, temos que $KMO = \frac{5,596}{7,965} = 0,703$. Portanto, $KMO > 0,5$ e obtemos aprovação no teste.

5.1.3.4 Cálculo e Avaliação das Componentes Principais

Tendo obtido aprovação nos testes, passa-se então para o cálculo completo e avaliação das componentes principais.

A Tabela 9 mostra os escores das cinco componentes selecionadas para cada uma das variáveis consideradas.

Tabela 9 – Escores das Componentes Principais de cada variável

Variável	Coeficientes de Ponderação (Escores)				
	Y1	Y2	Y3	Y4	Y5
Páginas de Relatório	0,414	-0,316	-0,114	-0,053	0,084
Páginas de Reivindicações	0,498	0,190	0,244	-0,006	-0,072
Páginas de Desenhos	0,332	-0,487	-0,224	-0,163	-0,052
Páginas de Subsídios	-0,059	-0,059	0,338	-0,668	0,653
Reivindicações Independentes	0,393	0,219	0,260	0,037	-0,053
Reivindicações Dependentes	0,457	0,250	0,247	-0,040	-0,129
Subclassess	0,032	-0,201	0,302	0,712	0,546
Ano de Depósito	-0,088	-0,456	0,500	0,044	-0,227
Inventores	0,217	-0,444	-0,273	-0,004	0,022
Prioridades	0,218	0,261	-0,468	0,114	0,436

Fonte: Elaboração Própria.

Na Tabela 9 foram hachurados na cor cinza os fatores considerados significativos para cada uma das variáveis (muito próximas ou acima de 0,4). Em uma primeira análise, nota-se que a componente Y1 é bastante destacada das demais. Além de explicar praticamente 30% de toda a variância, ela está associada às variáveis diretamente relacionadas ao volume de dados (páginas de relatório, de reivindicações e de desenhos, além de número de reivindicações independentes e dependentes). Tal fato também se encontra em consonância com a hipótese inicial 1, relativa às variáveis de volume. Já as componentes Y2 e Y3, embora possam ter alguma relação com volume de dados, representam basicamente as influências das variáveis, ano de depósito, número de inventores e prioridades. São componentes que aparentam estar associadas às estratégias de desenvolvimento e gestão dos depositantes, e maturidade da tecnologia envolvida. Por outro lado, as componentes Y4 e Y5 complementam as demais estando associadas às variáveis número de subclasses e páginas de subsídios. Tais componentes aparentam representar as influências específicas da área tecnológica dos pedidos de patente. Este resultado se encontra em consonância com a nossa hipótese inicial 2 relativa às variáveis com influência complementar ou indireta.

A Tabela 10 mostra a estrutura de dados das componentes principais selecionadas para cada um dos pedidos em análise.

Tabela 10 – Estrutura de dados das Componentes Principais selecionadas

Item	Examinador	Pedido	Subclasse Principal	Y5	Y4	Y3	Y2	Y1
1	1	PI0101329-7	H01C	0,26	-0,30	-1,95	0,04	-0,18
2	1	PI0111250-3	H01H	-0,20	-0,63	-0,48	1,31	2,22
3	1	PI0400832-4	H02K	0,12	-0,54	-1,18	0,83	0,60
4	1	PI0403974-2	H01B	0,19	-0,27	-0,99	1,65	-0,95
5	1	PI0107166-1	H01H	-0,29	-0,41	-0,71	0,94	-2,17
6	1	PI0303675-8	H02G	0,23	-0,45	-2,00	-0,68	0,37
.
.
.
814	11	PI0601239-6	H05B	-1,89407	-3,09	-5,54	-18,07	9,07

Fonte: Elaboração Própria.

A partir da obtenção das componentes principais de cada um dos pedidos de patente em análise, são calculadas as correlações entre as componentes e as variáveis originais, que são denominadas de Cargas das Componentes ou Cargas Fatoriais. Para cada uma das variáveis também são calculadas as somas dos quadrados das Cargas Fatoriais, denominadas de Comunalidades. As comunalidades expressam o quanto da variância de cada uma das

variáveis originais é explicável pelas componentes. Sendo assim, quanto maior a comunalidade de uma variável, maior será o poder das componentes em captar a influência da dita variável. Não há um consenso quanto ao valor adequado para as cargas fatoriais e comunalidades, mas autores e pesquisadores sugerem analisar somente fatores e variáveis com valores acima de 0,5. A Tabela 11 mostra os resultados obtidos para as cargas fatoriais e comunalidades.

Tabela 11 – Estrutura de dados das Componentes Principais selecionadas

Variável	Cargas Fatoriais (Correlações entre Fatores e Variáveis)					Comunalidades (Cargas Fatoriais ao quadrado)					Comunalidades Totais
	Y1	Y2	Y3	Y4	Y5	Y1	Y2	Y3	Y4	Y5	
Páginas de Relatório	0,711	-0,388	-0,122	-0,058	0,080	0,506	0,150	0,015	0,003	0,006	0,681
Páginas de Reivindicações	0,857	0,232	0,259	-0,013	-0,070	0,735	0,054	0,067	0,000	0,005	0,861
Páginas de Desenhos	0,572	-0,593	-0,235	-0,169	-0,051	0,327	0,352	0,055	0,029	0,003	0,765
Páginas de Subsídios	-0,099	-0,071	0,357	-0,677	0,628	0,010	0,005	0,128	0,458	0,394	0,995
Reivindicações Independentes	0,677	0,267	0,275	0,039	-0,052	0,458	0,071	0,076	0,002	0,003	0,610
Reivindicações Dependentes	0,787	0,309	0,259	-0,045	-0,123	0,619	0,095	0,067	0,002	0,015	0,799
Subclasses	0,052	-0,242	0,320	0,722	0,524	0,003	0,059	0,103	0,522	0,275	0,961
Ano de Depósito	-0,153	-0,556	0,527	0,043	-0,215	0,024	0,309	0,278	0,002	0,046	0,658
Inventores	0,374	-0,537	-0,295	-0,003	0,020	0,140	0,289	0,087	0,000	0,000	0,516
Prioridades	0,373	0,315	-0,495	0,111	0,421	0,139	0,099	0,245	0,012	0,177	0,673

Fonte: Elaboração Própria.

Na Tabela 11 foram hachurados na cor cinza os valores de cargas fatoriais acima de 0,5. Nota-se que a estrutura manteve-se semelhante a da Tabela 9 com os escores originais. Verifica-se ainda que todas as dez variáveis selecionadas apresentaram comunalidades acima de 0,5 e, sendo assim, todas elas são consideradas relevantes para o processo. Mais uma vez, os resultados sugerem que todas as variáveis, inclusive as que não são diretamente relacionadas ao volume de dados, tem alguma relevância na avaliação, ou seja, em consonância com a nossa hipótese inicial 2 relativa às variáveis com influência complementar ou indireta.

5.1.3.5 Avaliação dos Pedidos de Patente da Amostra de Teste Inicial

Foram calculados os índices IGC padronizados (média igual a zero e desvio padrão igual a um) de todos os pedidos de patente da amostra. A Figura 11 mostra a plotagem dos IGC padronizados (eixo y) relativos a cada um dos pedidos (eixo x).

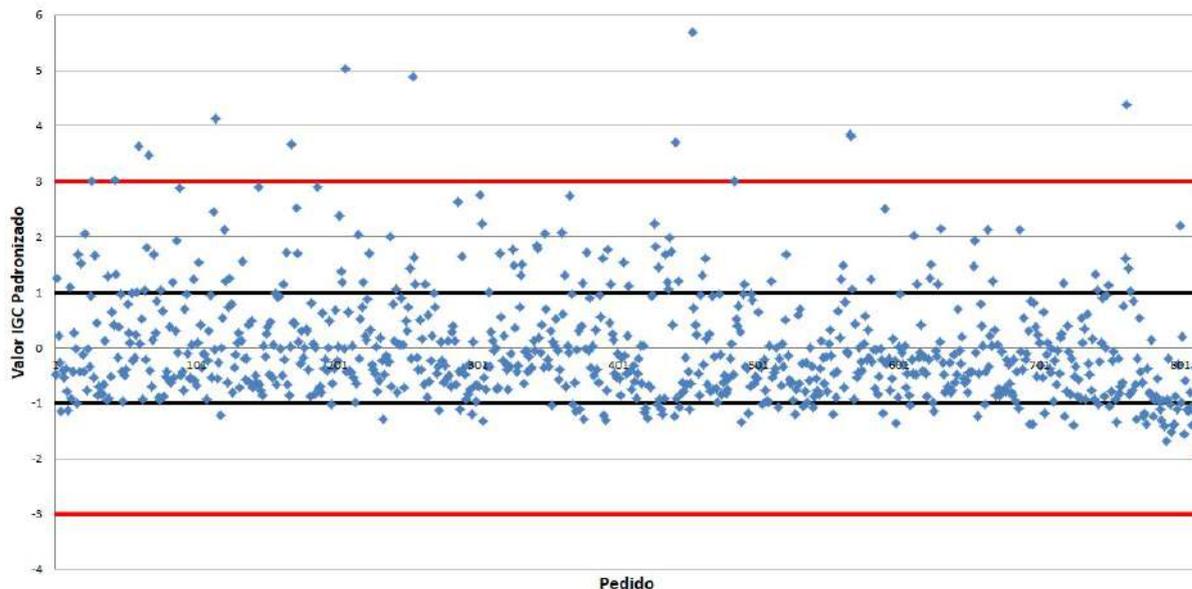


Figura 11 – IGC padronizado dos pedidos

Com base nos IGC padronizados, todos os pedidos da amostra foram então enquadrados conforme as faixas já definidas na Tabela 2, ou seja:

a) se IGC é maior que três ($IGC > 3$, ou seja, é maior do que a média mais três desvios padrão) então o pedido é muito pesado;

b) se IGC é maior do que um e menor ou igual a três ($1 < IGC \leq 3$, ou seja, está no intervalo entre média mais um e média mais três desvios padrão) então o pedido é pesado;

c) se IGC está no intervalo entre menos um e um ($-1 \leq IGC \leq 1$, ou seja, está no intervalo entre média mais ou menos um desvio padrão) então o pedido é moderado; e

d) se IGC está no intervalo entre menos três e menos um ($-3 \leq IGC < -1$, ou seja, está no intervalo entre média menos desvio padrão e média menos três desvios padrão) então o pedido é leve.

A Figura 12 mostra os enquadramentos encontrados.

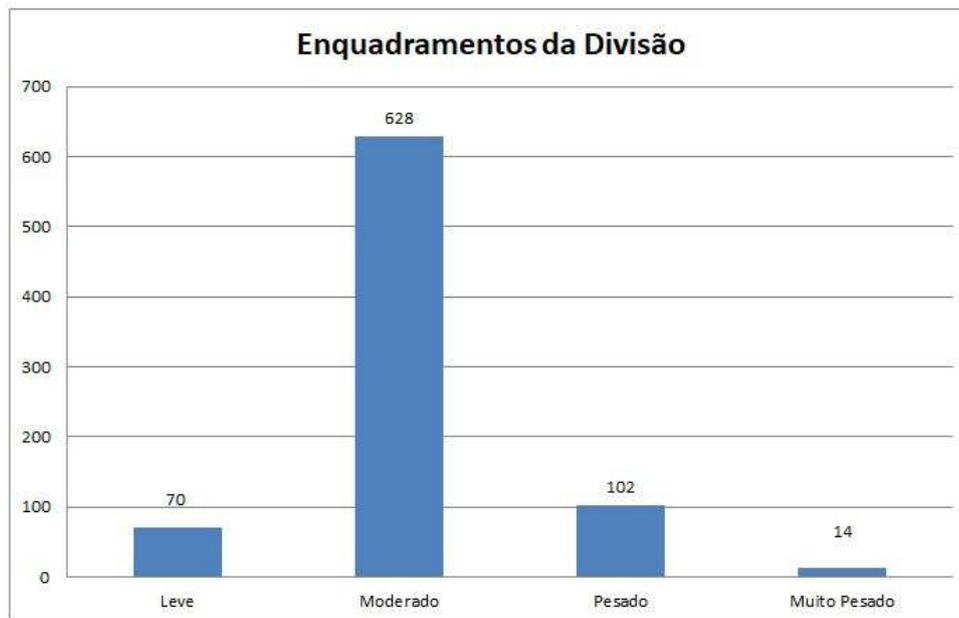


Figura 12 – Enquadramentos dos pedidos da amostra de teste inicial

Da análise das Figuras 11 e 12 verifica-se que nenhum dos pedidos de patente da amostra foi classificado como muito leve. A maioria dos pedidos foi classificada como moderado (628 ou 77,15%), em segundo lugar como pesados (102 ou 12,53%), seguidos pelos leves (70 ou 8,53%). Destaque-se que foram bem poucos os pedidos classificados como muito pesados (14 ou 1,79%).

Em uma primeira tentativa de verificar a coerência dos enquadramentos encontrados pelo IGC, estes foram comparados com os seguintes casos de enquadramentos mais intuitivos quanto ao volume de dados/complexidade:

- ✓ Caso 1: enquadramento dos pedidos considerando somente a variável número total de páginas;
- ✓ Caso 2: enquadramento dos pedidos considerando somente a variável número total de reivindicações.

Destaque-se que, para obter uma mesma referência quando da comparação, foram utilizadas as mesmas faixas de enquadramento para todos os casos a serem comparados, ou seja, nos dois novos casos os pedidos também foram enquadrados conforme as regras da Tabela 2, seguindo o mesmo padrão utilizado na construção do gráfico da Figura 12 acima. A Tabela 12 mostra os resultados obtidos. Nela é mostrado o número de pedidos enquadrados em cada uma das classes (Muito pesado, Pesado, Moderado e Leve) para cada um dos casos

(caso 1: enquadramento por total de páginas; caso 2: enquadramento por total de reivindicações; e enquadramento pelo IGC). É ainda mostrada a similaridade entre os enquadramentos pelo IGC pelas duas variáveis selecionadas. Esta similaridade é calculada como a porcentagem de enquadramentos idênticos entre os casos, ou seja, a relação entre o número de enquadramentos iguais e o total de pedidos enquadrados.

Tabela 12 – Comparação de enquadramentos pelo IGC, total de páginas e total de reivindicações

Classe	Número Total de Pedidos Enquadrados		
	Caso 1: Total de Páginas	Caso 2: Total Reivindicações	IGC
Muito Pesado	17	16	14
Pesado	61	86	102
Moderado	728	672	628
Leve	8	40	70
Similaridade IGC	76,66%	83,29%	

Fonte: Elaboração Própria.

Da análise da Tabela 12 verifica-se que há uma boa similaridade entre os enquadramentos obtidos por meio do IGC e por meio do número total de páginas (76,66%) e de reivindicações (83,29%). Este resultado era esperado dado que, a princípio, o volume de dados de um pedido de patente é um problema praticamente bidimensional, dependendo exatamente destas duas variáveis (total de páginas e reivindicações). Todavia, também era esperado que este resultado fosse alto, mas não igual a 100%, caso contrário não seria necessário o cálculo de um novo índice, dado que este não teria captado influências de outras variáveis. Pode-se notar ainda que a similaridade é mais alta com relação às reivindicações, variável esta que é das mais relevantes, senão a mais relevante de um pedido de patente.

Uma última comparação foi realizada entre os enquadramentos obtidos pelo IGC e considerando apenas a componente principal (Y1) dos pedidos. Mais uma vez, para termos uma mesma referência quando da comparação, foram utilizadas as mesmas faixas de enquadramento para ambos os casos a serem comparados, ou seja, os pedidos também foram enquadrados conforme as regras da Tabela 2, seguindo o mesmo padrão utilizado na Tabela 12. Os resultados podem ser vistos na Tabela 13, onde é mostrado o número de pedidos enquadrados em cada uma das classes (Muito pesado, Pesado, Moderado e Leve) para cada um dos casos comparados (Enquadrados pela componente principal e pelo IGC).

Tabela 13 – Comparação de enquadramentos pelo IGC e pela componente Y1

Classe	Número Total de Pedidos Enquadrados	
	Y1	IGC
Muito Pesado	16	14
Pesado	102	102
Moderado	623	628
Leve	73	70
Similaridade	86,61%	

Fonte: Elaboração Própria.

Os resultados mostrados na Tabela 13 evidenciam uma boa similaridade entre os enquadramentos pelo IGC e somente considerando a componente principal Y1. Este resultado mostra que de fato a primeira componente é aquela que melhor capta o volume de dados em geral dos pedidos.

De modo a evidenciar ainda mais a coerência dos enquadramentos por IGC, a tabela 14 a seguir mostra em detalhes os 14 pedidos enquadrados como muito pesados. São ainda mostradas as médias e medianas da amostra para cada uma das variáveis em análise.

Tabela 14 – Pedidos enquadrados como muito pesados por IGC

Pedidos Muito Pesados no Enquadramento por IGC											
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Reivindicações Independentes	Reivindicações Dependentes	Subclasses	Ano de Depósito	Inventores	Prioridades	Total de Páginas	Total de Reivindicações
27	51	12	7	6	46	2	1	2	1	70	52
43	68	13	17	8	26	2	1	2	1	98	34
60	36	20	6	5	55	1	2	3	2	62	60
67	71	11	21	4	4	3	2	1	7	103	8
115	13	26	5	7	56	1	2	5	1	44	63
169	27	12	13	8	73	1	3	3	2	52	81
207	68	22	14	9	74	1	3	1	1	104	83
255	12	23	4	8	88	1	3	3	1	39	96
442	62	19	15	7	56	1	4	4	1	96	63
454	33	24	19	10	76	2	4	3	2	76	86
484	24	13	6	4	70	1	4	1	2	43	74
566	14	14	7	11	40	3	5	3	1	35	51
567	32	15	9	6	61	2	5	2	1	60	67
763	62	20	19	10	28	3	7	6	2	101	38
Médias da Divisão	16,45	4,19	5,99	2,25	13,83	1,41	4,32	2,40	0,83	27,13	16,08
Medianas da Divisão	12,00	3,00	4,00	2,00	11,00	1,00	4,00	2,00	1,00	20,00	13,00

Fonte: Elaboração Própria.

Da análise da tabela 14, verifica-se que todos os pedidos enquadrados como muito pesados pelo IGC apresentam as cinco variáveis de volume com valores quase sempre bem acima da média e mediana da divisão. Além disso, todos os pedidos apresentam número total de páginas e número total de reivindicações também bastante acima da média e mediana da divisão. Cabe um destaque para o pedido de número 67 que, embora tenha somente 8 reivindicações, apresenta um total de 103 páginas (muito acima da média e mediana) e ainda

um total de 11 páginas de reivindicações (muito acima da média e mediana), ou seja, é um pedido muito grande e com reivindicações também muito grandes. Sendo assim, os resultados sugerem que os enquadramentos estão de fato em consonância com a voluminosidade dos pedidos avaliados.

5.1.4 Análise de Subclasses da CIP

5.1.4.1 ZAE dos Examinadores

Na amostra de teste inicial foi encontrado um número total de 95 subclasses principais ao longo de todos os 814 pedidos de patentes analisados. Todavia, ao se considerar apenas as ZAE dos examinadores (subclasses com 5% ou mais de pedidos de cada examinador), se verifica que 25 subclasses principais foram responsáveis por enquadrar 636 pedidos, ou seja, cerca de 80% do total avaliado. Cabe destacar que, dado que 3 destas 25 subclasses tiveram ocorrências muito baixas, foram utilizadas 22 subclasses no conjunto de avaliação de interesse, equivalendo a 619 pedidos (76% do total).

A Tabela 15 mostra as zonas de atuação de cada um dos 11 examinadores por subclasses da CIP. Destaca-se que a área cinza corresponde às Zonas de Atuação Específicas (ZAE) dos examinadores, enquanto a área branca corresponde às subclasses que, embora não pertençam à ZAE do examinador, são pertencentes às ZAE de algum dos demais examinadores avaliados.

Tabela 15 – Zonas de Atuação dos Examinadores por Subclasses da CIP

Zonas de Atuação dos Examinadores por Subclasses																					
1		2		3		4		5		6		7		8		9		10		11	
Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%
H01H	63,08%	H01R	33,33%	H05B	18,42%	H02K	32,65%	H02J	19,54%	H02P	12,26%	H02K	56,36%	B66B	31,58%	G07F	17,09%	H01R	52,17%	B66B	12,09%
H02G	15,38%	H01H	23,33%	G01R	18,42%	H01M	14,29%	H02M	16,09%	B60L	9,43%	H01B	9,09%	B50T	14,04%	G08B	11,11%	H02G	24,64%	G01R	10,99%
H01B	6,15%	H01B	7,78%	H02P	15,79%	B60L	6,12%	G01R	13,79%	H02J	6,60%	H01R	7,27%	B60K	5,26%	B66B	7,69%	H02B	4,35%	H02B	8,79%
H01F	3,08%	G01R	6,67%	H01F	10,53%	G01R	4,08%	F02D	10,34%	F02D	6,60%	H01F	5,45%	F02D	5,26%	H05B	5,13%	H01H	2,90%	H01H	6,59%
H02K	1,54%	H05B	4,44%	H02H	7,89%	H02J	2,04%	B60K	4,60%	H01R	5,66%	F21V	3,64%	H02P	3,51%	F21V	5,13%	H02H	1,45%	H05B	4,40%
H02P	1,54%	H02H	3,33%	H01M	7,89%	F21V	2,04%	H02H	4,60%	B60K	3,77%	G01R	1,82%	H02G	1,75%	G01R	3,42%	H02K	1,45%	F21V	3,30%
H02B	1,54%	H02J	3,33%	H01H	5,26%	H02H	2,04%	H01M	4,60%	H02K	3,77%	H02G	1,82%	H02H	1,75%	H02H	2,56%			B60L	2,20%
		H02K	2,22%	H02G	2,63%	H02P	2,04%	H02G	1,15%	H01F	3,77%	H02B	1,82%			H01H	2,56%			H01R	1,10%
		H02G	2,22%	B60K	2,63%	H01H	2,04%	H01F	1,15%	H01M	3,77%	H01M	1,82%			H02J	1,71%			H01M	1,10%
				H02K	2,63%	H02B	2,04%	B60L	1,15%	H05B	3,77%					H01M	1,71%			H02J	1,10%
						H02G	2,04%	H01H	1,15%	H02H	2,83%					H02K	0,85%			H02H	1,10%
										G01R	1,89%					H01F	0,85%			B50T	1,10%
										F21V	0,94%					B60L	0,85%			H02P	1,10%
										H01H	0,94%										

Fonte: Elaboração Própria

Da análise da Tabela 15 se verifica que as ZAE de cada um dos examinadores são bem variáveis. Enquanto o examinador 10 apresenta uma concentração de exame em somente duas subclasses principais, os examinadores 1 e 4 apresentam três subclasses, os examinadores 2, 5, 7, 8 e 11 apresentam quatro subclasses, os examinadores 6 e 9 apresentam cinco subclasses e o examinador 3 apresenta 7 subclasses. Tal fato mostra ser evidente uma distribuição de carga atual utilizando como base a especialidade/áreas de conhecimento dos examinadores.

5.1.4.2 Avaliação das Subclasses da CIP das ZAE

Para identificar o comportamento dos dados dentro de cada uma das subclasses principais encontradas, foram obtidas as médias das principais variáveis dos pedidos por subclasse. Obtidas as médias, foram calculados os desvios padrão e coeficientes de variação. A Tabela 16 mostra estes resultados.

Tabela 16 – Médias, Desvios Padrão e Coeficientes de Variação das variáveis por Subclasse e por Pedido

Subclasses	Médias por Pedido Examinado								Nr. Subclasses
	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações		
B60K	16,75	4,92	5,92	27,58	2,25	15,08	17,33	2,25	
B60L	17,47	6,06	7,41	30,94	3,12	16,82	19,94	1,65	
B60T	37,08	5,75	6,42	49,25	2,50	16,75	19,25	1,67	
B66R	23,84	3,55	8,76	36,16	2,24	10,92	13,16	1,05	
F02D	18,78	3,65	6,22	29,70	2,26	9,48	11,74	1,48	
F21V	10,69	3,62	3,15	17,46	2,77	11,00	13,77	1,77	
G01R	16,93	3,70	5,39	26,02	2,32	13,05	15,36	1,50	
G07F	9,95	3,20	4,10	17,25	1,80	10,75	12,55	1,41	
G08B	18,92	5,69	6,15	30,77	3,08	17,54	20,62	1,62	
H01B	22,75	5,63	7,38	35,75	2,88	23,75	26,63	1,13	
H01F	18,13	3,07	6,80	28,00	1,80	8,33	10,13	1,60	
H01H	10,71	3,43	5,03	19,17	1,78	12,14	13,92	1,17	
H01M	17,15	4,50	5,90	29,20	2,20	15,80	18,00	1,20	
H01R	13,30	4,49	6,21	27,03	2,52	15,92	18,25	1,14	
H02B	8,92	2,54	3,85	15,31	1,46	9,15	10,62	1,38	
H02G	14,55	3,67	6,79	25,64	1,48	13,73	15,21	1,42	
H02H	16,20	4,95	4,60	25,75	1,95	15,00	16,95	1,80	
H02I	19,71	5,39	7,71	32,81	2,77	16,97	19,74	1,81	
H02K	14,81	3,47	6,21	24,49	2,19	11,42	13,61	1,35	
H02M	15,05	3,63	5,37	24,05	2,16	11,89	14,05	1,32	
H02P	18,72	4,80	4,60	28,12	2,52	13,88	16,40	1,36	
H05B	12,13	4,29	6,71	23,13	2,38	11,08	13,46	1,46	
Média Divisão	16,45	4,19	5,99	27,13	2,25	13,83	16,08	1,41	
Desvio Padrão	5,97	1,00	1,35	7,39	0,46	3,57	3,87	0,28	
Coef. Variação	36,31%	23,92%	22,61%	27,25%	20,44%	25,80%	24,04%	19,91%	

Fonte: Elaboração Própria

Da análise da Tabela 16 se verifica que, embora a dispersão de todas as variáveis dos pedidos médios equivalentes por subclasses tenha sido menor do que a dispersão do conjunto total de pedidos (Ver Tabela 4), essa dispersão ainda é elevada, sempre com coeficiente de variação acima de 20% e, em alguns casos, beirando 30% ou 40%. Sendo assim, percebe-se que há uma grande diferença no perfil do volume de dados de cada uma das subclasses. Vale ressaltar que, analisando cada classe individualmente, os coeficientes de variação são

elevados, da ordem de 30% até 100%. Ou seja, há uma variabilidade enorme de pedidos dentro de cada uma das subclasses. A Tabela 17 mostra estes resultados.

Tabela 17 – Coeficientes de Variação das Variáveis de cada Subclasse

Subclasses	Médias por Pedido Examinado - Coeficientes de Variação						
	Páginas de Relatório	Páginas de Reivindicações	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações	Nr. Subclasses
B60K	62,93%	67,43%	55,64%	54,02%	55,21%	50,44%	57,25%
B60L	65,30%	72,29%	67,81%	80,89%	71,24%	64,62%	52,32%
B60T	130,79%	69,22%	117,58%	84,42%	76,67%	74,10%	59,08%
B66B	96,27%	59,15%	92,02%	63,59%	69,52%	60,43%	21,50%
F02D	71,86%	86,83%	66,33%	110,95%	93,77%	92,40%	49,41%
F21V	67,18%	89,30%	63,07%	88,64%	104,58%	86,09%	57,24%
G01R	76,17%	55,41%	64,30%	75,45%	74,47%	70,15%	46,59%
G07F	57,46%	60,49%	48,57%	68,87%	110,80%	96,08%	47,52%
G08B	55,11%	56,65%	42,83%	58,53%	62,27%	60,21%	47,54%
H01B	79,47%	67,42%	67,40%	70,57%	72,69%	70,58%	30,36%
H01F	91,64%	82,19%	86,72%	56,34%	61,30%	53,26%	39,53%
H01H	54,63%	82,69%	51,32%	60,91%	84,68%	79,32%	37,76%
H01M	51,89%	65,49%	45,43%	61,87%	76,30%	72,21%	34,20%
H01R	56,81%	76,38%	104,99%	81,51%	81,81%	79,25%	36,78%
H02B	32,18%	44,38%	31,93%	53,11%	44,79%	44,28%	46,98%
H02G	54,08%	50,87%	49,17%	47,96%	77,19%	72,10%	39,37%
H02H	85,59%	78,51%	74,92%	71,51%	87,30%	81,76%	34,20%
H02J	79,29%	77,68%	76,06%	71,70%	87,73%	79,67%	61,34%
H02K	70,85%	70,12%	63,88%	69,71%	69,26%	65,27%	40,77%
H02M	49,54%	70,03%	51,38%	69,53%	88,97%	83,95%	36,30%
H02P	90,33%	89,20%	80,49%	82,70%	103,81%	96,10%	51,47%
H05B	67,85%	82,30%	71,73%	64,29%	84,80%	75,17%	45,12%

Fonte: Elaboração Própria

Destaque-se que esta variabilidade interna das subclasses é uma característica interessante, pois evidencia uma tendência de eficiência para distribuição pelas ZAE.

A Tabela 18 mostra os enquadramentos dos pedidos por IGC padronizado dentro das subclasses das ZAE.

Tabela 18 – Enquadramentos dos pedidos por IGC por subclasse

Subclasse	Número de Pedidos					% de Pedidos			
	Muito Pesado	Pesado	Moderado	Leve	Total	Muito Pesado	Pesado	Moderado	Leve
B60K	0	4	8	0	12	0,00%	33,33%	66,67%	0,00%
B60L	0	5	11	1	17	0,00%	29,41%	64,71%	5,88%
B60T	1	3	8	0	12	8,33%	25,00%	66,67%	0,00%
B66B	0	1	36	1	38	0,00%	2,63%	94,74%	2,63%
F02D	1	1	21	0	23	4,35%	4,35%	91,30%	0,00%
F21V	0	1	9	3	13	0,00%	7,69%	69,23%	23,08%
G01R	0	3	37	4	44	0,00%	6,82%	84,09%	9,09%
G07F	0	2	11	7	20	0,00%	10,00%	55,00%	35,00%
G08B	0	4	8	1	13	0,00%	30,77%	61,54%	7,69%
H01B	0	6	8	2	16	0,00%	37,50%	50,00%	12,50%
H01F	1	0	13	1	15	6,67%	0,00%	86,67%	6,67%
H01H	1	4	66	6	77	1,30%	5,19%	85,71%	7,79%
H01M	0	4	14	2	20	0,00%	20,00%	70,00%	10,00%
H01R	2	8	62	5	77	2,60%	10,39%	80,52%	6,49%
H02B	0	0	13	0	13	0,00%	0,00%	100,00%	0,00%
H02G	0	3	27	3	33	0,00%	9,09%	81,82%	9,09%
H02H	1	4	14	1	20	5,00%	20,00%	70,00%	5,00%
H02J	2	6	20	3	31	6,45%	19,35%	64,52%	9,68%
H02K	0	4	51	2	57	0,00%	7,02%	89,47%	3,51%
H02M	0	1	16	2	19	0,00%	5,26%	84,21%	10,53%
H02P	1	5	17	2	25	4,00%	20,00%	68,00%	8,00%
H05B	0	3	19	2	24	0,00%	12,50%	79,17%	8,33%

Fonte: Elaboração Própria

Os dados da Tabela 18 mostram que, embora existam algumas variações, em todas as subclasses há uma predominância de pedidos enquadrados como moderados. Cabe destacar que não houve uma subclasse com grande concentração de pedidos enquadrados como muito pesados, pesados ou leves. Portanto, ao realizar uma distribuição considerando as ZAE (onde já foi visto que há grandes dispersões) em conjunto com os enquadramentos em classes (mostram grande concentração nos moderados), evidencia-se uma maior tendência de equilíbrio na distribuição do que considerando apenas as ZAE.

5.1.4.3 Redistribuição dos Pedidos

5.1.4.3.1 Nova configuração das ZAE dos examinadores

A Tabela 19 mostra a nova configuração das ZAE dos examinadores após ser realizada uma distribuição com a nova lógica desenvolvida (conforme itens 4.3.1 e 4.3.2). Cabe destacar que, na nova distribuição, o número de pedidos distribuídos para cada examinador permaneceu o mesmo que na distribuição original.

Tabela 19 – ZAE dos examinadores após a nova distribuição

Novas Zonas de Atuação dos Examinadores por Subclasses																							
1		2		3		4		5		6		7		8		9		10		11			
Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%		
H01H	62,50%	H01R	40,91%	H05B	55,26%	H02K	70,00%	H02J	31,03%	H02P	22,43%	H02K	52,73%	B66B	34,48%	G07F	17,09%	H01R	59,42%	B66B	19,78%		
H02G	17,19%	H01H	25,00%	H02H	39,47%	H01M	7,50%	H02M	21,84%	B60L	12,15%	H01B	3,64%	B60T	20,69%	G08B	11,11%	H02G	31,88%	G01R	27,47%		
H01B	10,94%	H01B	7,95%	H01H	2,63%	B60L	10,00%	G01R	19,54%	H02J	3,74%	H01F	27,27%	B60K	20,69%	H05B	2,56%					H02B	14,29%
		G01R	2,27%					F02D	12,64%	F02D	8,41%			F02D	5,17%	F21V	11,11%					H01H	15,38%

Fonte: Elaboração Própria

A Tabela 20 mostra uma comparação das porcentagens de pedidos dentro das ZAE em relação ao total de pedidos distribuídos para cada examinador nas distribuições nova e original.

Tabela 20 – Porcentagem de pedidos dentro das ZAE antes e após a nova distribuição

Examinador	1	2	3	4	5	6	7	8	9	10	11
Nova ZAE	90,63%	76,14%	97,37%	87,50%	85,06%	46,73%	83,64%	81,03%	41,88%	91,30%	76,92%
ZAE original	84,62%	71,11%	84,21%	53,06%	59,77%	40,57%	78,18%	56,14%	46,15%	76,81%	38,46%

Fonte: Elaboração Própria

Da análise das Tabelas 19 e 20 verifica-se que nenhuma nova subclasse, além das já integrantes, foi incluída nas ZAE dos examinadores, o que era um objetivo. Pelo contrário, algumas ZAE foram até mesmo enxugadas e concentradas. Ressalte-se que no caso de dez dos onze examinadores houve um expressivo aumento do número de pedidos distribuídos a eles e pertencentes às suas próprias ZAE. Sendo assim, estes resultados sugerem que esta nova configuração contribui para que os examinadores trabalhem dentro das suas áreas de formação e conhecimento específicas e, conseqüentemente, para a sua eficiência e motivação.

5.1.4.3.2 Nova configuração dos pedidos redistribuídos

A Tabela 21 mostra as medianas das variáveis dos pedidos redistribuídos com os dados separados por examinador. Nesta tabela foram hachurados: na cor cinza claro as variáveis dos pedidos de examinadores com mediana abaixo da mediana geral dos dados; e na cor cinza escuro as variáveis dos pedidos dos examinadores com mediana acima da mediana geral dos dados.

Tabela 21 – Medianas da Amostra de Teste Inicial redistribuída separadas por examinador

	Examinador	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Páginas de Subsídios	Reivindicações Independentes	Reivindicações Dependentes	Subclasses	Ano Depósito	Inventores	Prioridades
Medianas	1	12,50	3,00	4,00	0,00	1,00	9,50	1,00	3,00	2,00	1,00
	2	12,00	3,00	5,00	0,00	2,00	11,00	1,00	4,00	2,00	1,00
	3	11,00	3,00	4,00	0,00	2,00	9,50	1,00	4,00	2,00	1,00
	4	10,00	3,00	4,00	0,00	2,00	11,00	1,00	4,00	1,50	1,00
	5	14,00	3,00	5,00	0,00	2,00	11,00	1,00	5,00	2,00	1,00
	6	13,00	3,00	4,00	0,00	2,00	11,00	1,00	4,00	2,00	1,00
	7	13,50	3,00	5,00	0,00	2,00	10,00	1,00	4,00	2,00	1,00
	8	14,00	3,00	4,00	0,00	2,00	11,00	1,00	4,00	2,00	1,00
	9	12,00	3,00	4,00	0,00	2,00	11,00	1,00	4,00	2,00	1,00
	10	11,00	3,00	5,00	0,00	2,00	13,00	1,00	4,00	2,00	1,00
	11	11,00	3,00	4,00	0,00	2,00	9,00	1,00	5,00	2,00	1,00
	Geral	12,00	3,00	4,00	0,00	2,00	11,00	1,00	4,00	2,00	1,00

Fonte: Elaboração Própria.

Da análise da Tabela 21, verifica-se que ainda há algumas variáveis de pedidos dos examinadores com mediana tanto abaixo quanto acima da mediana geral dos dados da divisão. Todavia, nota-se que agora estes desvios foram reduzidos (quando comparados com a distribuição original – vide tabela 6), sendo que algumas variáveis essenciais visivelmente apresentaram melhora, tais como: páginas de reivindicações e reivindicações independentes. O examinador de número 5 ainda parece ter uma tendência a estar levemente sobrecarregado em relação aos demais, enquanto os examinadores de números 7 e 11 parecem ter entrado em uma posição moderada junto aos demais.

5.1.4.3.3 Índice de Balanceamento da Divisão (IBD) antes e após a redistribuição

Para complementar o ciclo da metodologia e realizar uma última medida de comparação entre as distribuições, foram calculados os índices IBD da distribuição original da amostra e da sua redistribuição. Para realização dos cálculos foram utilizadas as medianas constantes da tabela 6 (distribuição original) e da tabela 21 (redistribuição).

Por meio da equação (16), foram então obtidos IBD igual a 0,86 para o caso original e IBD igual a 0,88 para a redistribuição, ou seja, houve um aumento do IBD com nova distribuição que, em conjunto com a análise da tabela 21, reforça que as medianas dos pedidos dos examinadores após a redistribuição se aproximam mais da mediana geral da divisão. Tal fato corrobora que com a nova distribuição temos uma tendência de maior equilíbrio quanto ao volume de dados/complexidade dos pedidos distribuídos aos examinadores.

Como complexidade é algo relativo, para ter-se mais um indicativo se esse volume/complexidade de fato capta o tempo/esforço de exame, faz-se a seguir uma análise de sensibilidade do modelo de modo a verificar as correlações das variáveis e índices com o tempo de exame.

5.2 AMOSTRA PADRÃO COM TEMPO

5.2.1 Composição da Amostra e Variáveis de Interesse

De maneira análoga ao procedimento realizado para a amostra de teste inicial, foram obtidos dados de pedidos de patente da área tecnológica referente à engenharia elétrica, mais especificamente a Divisão de Eletricidade da DIRPA. Foram seguidas as etapas da metodologia proposta no que concerne à seleção do marco temporal e demais premissas definidas quanto ao arcabouço documental a ser analisado. Todavia, como neste caso o objetivo era obter uma amostra padrão incluindo o tempo dispendido no exame para servir de referência, todos os dados desta amostra foram a partir do formulário de tempo de exame preenchido por um único examinador. A escolha de um único examinador teve como intuito evitar fontes de erro exógenas adicionais para o modelo, principalmente devido a discrepâncias nos dados por possíveis comportamentos peculiares de examinadores diferentes, mas também pelo examinador escolhido ser de fato reconhecido pelos pares como referência na divisão.

Para cada um dos pedidos de patente do referido examinador foram obtidas todas as variáveis de interesse, perfazendo um total de cem (100) pedidos de patente, todos os primeiros exames já publicados, a serem avaliados e compoendo a amostra padrão. Os dados foram coletados entre janeiro de 2020 e fevereiro de 2021, sendo que todos os pedidos da amostra são do tipo 6.21, ou seja, com aproveitamento de buscas de escritórios internacionais.

Como a amostra padrão com tempo foi obtida para pedidos de patente enquadrados no plano de combate ao *backlog*, das dez possíveis variáveis analisadas na amostra inicial, a única que não foi considerada neste caso foi o número de páginas de subsídios, dado que não há nenhum pedido com subsídios presente na amostra.

As Tabelas 22 e 23 mostram os formulários de tempo de exame preenchidos pelo examinador.

Tabela 22 – Formulário de tempo de exame (parte 1) preenchido

Formulário para Preenchimento do Tempo despendido para cada uma das Etapas do Exame Técnico														
		Etapa 1 - Leitura e entendimento do Pedido (Preenchimento Quadros 1 a 3)				Etapa 2 - Busca de Anterioridades (Preenchimento Quadro 4)			Etapa 3 - Análise de Patenteabilidade (Preenchimento Quadro 5)					
Número Pedido	Tipo	Relatório Descritivo (minutos)	Quadro Reivindicatório (minutos)	Análise Desenhos (minutos)	Total Etapa 1 (horas)	Formulação da Estratégia (minutos)	Execução da Estratégia (minutos)	Total Etapa 2 (horas)	Análise características Reivindicações (minutos)	Análise características anterioridades (minutos)	Comparação e Justificativas Técnicas (minutos)	Total Etapa 3 (horas)	Tempo Total Pedido (horas)	
1	BR112012007070-3	6.21	60	30	45	2,25	0	0	0	90	200	100	6,50	8,75
2	BR102012015806-0	6.21	120	60	80	4,33	0	0	0	160	180	100	7,33	11,67
3	BR112013020584-9	6.21	110	140	40	4,83	0	0	0	200	210	110	8,67	13,50
4	BR112014009426-8	6.21	110	100	90	5,00	0	0	0	190	210	110	8,50	13,50
5	PI1105866-8	6.21	80	95	50	3,75	0	0	0	190	220	130	9,00	12,75
6	PI0924510-3	6.21	360	180	120	11,00	0	0	0	280	260	160	11,67	22,67
.
.
.
100	BR112014015918-1	6.21	130	180	90	6,67	0	0	0	280	260	160	11,67	18,33

Fonte: Elaboração Própria.

Tabela 23 – Formulário de tempo de exame (parte 2) preenchido

		Avaliação do grau de esforço/dificuldade De 0 a 10, preencher que nota daria para o pedido (Quanto mais próximo de 10 mais difícil)		
	Número Pedido	Tempo Total Pedido (horas)	Nota (0 a 10)	Espaço para Justificativa (caso ache necessário)
1	BR112012007070-3	8,75	5	Etapas: 6.21 - 9.1. Primeiro Exame
2	BR102012015806-0	11,67	4	Etapas: 6.21 - 9.1. Primeiro Exame
3	BR112013020584-9	13,50	6	Etapas: 6.21 - 6.1. Primeiro Exame
4	BR112014009426-8	13,50	7	Etapas: 6.21 - 6.1. Primeiro Exame
5	PI1105866-8	12,75	7	Etapas: 6.21 - 7.1. Primeiro Exame
6	PI0924510-3	22,67	10	Etapas: 6.21 - 9.1. Primeiro Exame
.
.
.
100	BR112014015918-1	18,33	9	Etapas: 6.21 - 6.1. Primeiro Exame

Fonte: Elaboração Própria.

A Tabela 24 mostra a estrutura dos dados da amostra padrão.

Tabela 24 – Estrutura da Amostra Padrão

Item	Pedido	Subclasse Principal	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Reivindicações Independentes	Reivindicações Dependentes	Total de Subclasses	Ano de Depósito	Inventores	Prioridades	Tempo
1	BR112012018158-0	F02M	6	4	3	2	13	4	11	1	1	10,67
2	BR112012018368-0	F02M	6	3	3	3	6	2	11	2	1	12,75
3	BR112013032420-1	H01H	10	1	1	1	0	2	12	2	1	7,92
4	BR112012018119-0	G01P	11	4	2	1	6	3	10	1	1	10,08
.
.
.
100	BR112014013523-1	B60W	14	3	5	1	8	3	12	4	1	14,33

Fonte: Elaboração Própria.

A Tabela 25 mostra, para o conjunto total de 100 pedidos, um quadro geral das principais variáveis obtidas e os resultados dos cálculos de estatísticas básicas tais como: média, desvio padrão, mediana, moda, mínimo, máximo e coeficiente de variação.

Tabela 25 – Estatísticas básicas da Amostra Padrão

	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Reivindicações Independentes	Reivindicações Dependentes	Total de Subclasses	Ano de Depósito	Inventores	Prioridades	Tempo
Média	21,63	3,89	5,88	2,01	9,62	1,52	11,68	2,25	0,93	13,69
Desvio Padrão	13,82	1,85	4,14	0,96	5,50	0,92	0,72	1,49	0,56	2,96
Mediana	19,00	4,00	5,00	2,00	8,50	1,00	12,00	2,00	1,00	13,50
Moda	16	3	4	2	4	1	12	1	1	12,75
Mínimo	4	1	0	1	0	1	9	1	0	7,92
Máximo	88	11	20	5	30	6	12	7	4	22,67
Coeficiente de Variação	63,90%	47,63%	70,36%	47,70%	57,22%	60,22%	6,19%	66,07%	59,69%	21,61%

Fonte: Elaboração Própria.

Da análise da Tabela 25, verifica-se que, para quase todas as variáveis do pedido a serem consideradas, os desvios padrão (e variância, pois é o quadrado do desvio) e coeficientes de variação (relação entre o desvio padrão e a média) ainda são altos, embora menores do que os da amostra de teste inicial. A exceção é o ano de depósito com um coeficiente de variação bem baixo (6,19%) e moda igual a doze, evidenciando que a distribuição para o examinador ficou bastante concentrada em um único ano de depósito, o de 2012. Além do ano de depósito, o tempo também apresentou um coeficiente de variação relativamente baixo (21,61%) em comparação com as demais variáveis, mas ainda acima de 15%, valor limite para que se considere verdadeiramente baixo. Adicionalmente, modas, médias e medianas das variáveis são distintas entre si, sendo que os valores mínimos e máximos são bastante diferentes. Tais fatos mais uma vez sugerem que alta amplitude e boa dispersão são características intrínsecas dos pedidos analisados, ou seja, há uma variabilidade razoável entre eles mesmo considerando uma amostra menor e com pedidos de mesmo tipo (6.21 com aproveitamento de buscas). Neste contexto, mais uma vez se verifica que uma análise que considere essa dispersão seria interessante.

5.2.2 Simulações e Testes: análise de sensibilidade do modelo

Para a realização da análise de sensibilidade do modelo, foram realizadas dezenas de simulações considerando desde o caso mais completo, com nove variáveis, até casos mais simples com três variáveis. Para todos os casos foram executadas simulações a cada dez pedidos constantes da amostra, ou seja, para cada conjunto de casos com três até nove variáveis, foram feitos testes considerando 10, 20, 30, ... , 90 e 100 pedidos de patente da amostra padrão. O número mínimo de dez pedidos para testes foi escolhido, pois é indicado que a amostra tenha no mínimo uma população maior do que o número de variáveis para aplicação do método da ACP, sendo que quanto maior a amostra, teoricamente melhor para o modelo.

Ao executar as simulações foram verificadas as correlações de todas as variáveis e dos índices IGC com o tempo, sendo que o IGC foi calculado tanto pelo critério de 70% da variância, sendo denominado $IGC_{70\%}$, quanto pelo critério de autovalores maiores do que um, sendo denominado $IGC_{\lambda>1}$. Vale ressaltar que, quando $IGC_{70\%}$ e $IGC_{\lambda>1}$ forem iguais, denominaremos simplesmente IGC. Por fim, foi ainda calculado um IGC somente relativo a componente principal (Y1) dos casos, a componente mais significativa em variância, sendo denominado IGC_{Y1} .

Após a realização de todas as simulações e de posse de uma gama enorme de resultados para dezenas de casos, foram selecionados os casos de maior relevância e de maior interesse em termos variáveis analisadas e suas correlações com o tempo.

5.2.2.1 Casos de estudo: aplicação da ACP e análise de correlações

5.2.2.1.1 Caso 1: Caso 9 Var

O primeiro caso a ser analisado foi o caso mais completo com nove variáveis, denominado Caso 9 Var, incluindo as seguintes variáveis:

- ✓ número de páginas do relatório descritivo (variável 1);
- ✓ número de páginas de reivindicações (variável 2);
- ✓ número de páginas de desenhos (variável 3);
- ✓ número de reivindicações independentes (variável 4);
- ✓ número de reivindicações dependentes (variável 5);
- ✓ número de subclasses da CIP (variável 6);
- ✓ ano de depósito (variável 7);

- ✓ número de inventores (variável 8); e
- ✓ número de prioridades (variável 9).

5.2.2.1.1.1 Cálculo da Matriz de Correlação e Seleção das Componentes Principais

De maneira análoga aos passos executados quando da análise da amostra de teste inicial, após a padronização dos dados, com a aplicação das equações (5) ou (6), são então determinadas as matrizes S e R das variáveis padronizadas e, como são nove as variáveis utilizadas, ambas as matrizes agora são iguais e quadradas de dimensão 9 x 9. Como foram executados dez casos diferentes (com 10 até 100 pedidos), a título de ilustração é mostrada a seguir a matriz S=R relativa à amostra completa, com 100 pedidos.

$$S = R = \begin{bmatrix} 1 & 0,52 & 0,63 & 0,16 & 0,22 & -0,05 & -0,13 & 0,16 & 0,25 \\ 0,52 & 1 & 0,29 & 0,34 & 0,70 & 0 & -0,18 & 0,07 & 0,12 \\ 0,63 & 0,29 & 1 & 0 & -0,01 & -0,01 & -0,08 & 0,18 & 0,03 \\ 0,16 & 0,34 & 0 & 1 & 0,37 & -0,03 & -0,01 & -0,12 & 0,21 \\ 0,22 & 0,70 & -0,01 & 0,37 & 1 & 0,02 & -0,16 & -0,08 & 0,08 \\ -0,05 & 0 & -0,01 & -0,03 & 0,02 & 1 & -0,11 & 0,07 & 0,17 \\ -0,13 & -0,18 & -0,08 & -0,01 & -0,16 & -0,11 & 1 & 0,12 & 0,27 \\ 0,16 & 0,07 & 0,18 & -0,12 & -0,08 & 0,07 & 0,12 & 1 & -0,03 \\ 0,25 & 0,12 & 0,03 & 0,21 & 0,08 & 0,17 & 0,27 & -0,03 & 1 \end{bmatrix}$$

Tendo obtido as matrizes S e R foram então aplicadas as equações (7) a (9) para determinação dos autovalores ($\lambda_1, \lambda_2, \dots, \lambda_9$) e da matriz V de autovetores associados (scores das componentes).

De posse dos autovalores e autovetores, são então aplicadas as equações (10) e (11) para determinação das componentes principais e das suas respectivas contribuições de variâncias. É ainda calculada a contribuição acumulada das componentes (% de variância acumulada). A Tabela 26 mostra estes resultados.

Tabela 26 – Autovalores e Contribuições de Variância das Componentes Principais

Componente Principal	Autovalores	Variância	Variância Acumulada
Y1	2,47	27,42%	27,42%
Y2	1,51	16,74%	44,17%
Y3	1,33	14,75%	58,91%
Y4	1,09	12,09%	71,00%
Y5	0,92	10,24%	81,24%
Y6	0,66	7,28%	88,52%
Y7	0,54	5,96%	94,49%
Y8	0,28	3,08%	97,57%
Y9	0,22	2,43%	100,00%

Fonte: Elaboração Própria.

Da análise da Tabela 26, tanto no caso de utilização do critério considerando autovalores maiores do que um, quanto pela variância acumulada acima de 70%, seriam selecionadas as 4 (quatro) primeiras componentes para a determinação do IGC. Esse mesmo fenômeno foi observado para todas simulações com variações das amostras, de 10 até 100 pedidos.

Os testes de Bartlett e KMO foram aplicados para cada um dos casos. Quanto ao Teste de Bartlett, com exceção do caso 9 var com somente 10 pedidos, todos os demais foram aprovados. E no teste KMO, todos os casos acima de 40 pedidos foram aprovados.

As Figuras 13 e 14 mostram, respectivamente, os autovalores e as variâncias acumuladas para cada um dos casos 9 var, com amostras de 10 até 100 pedidos.

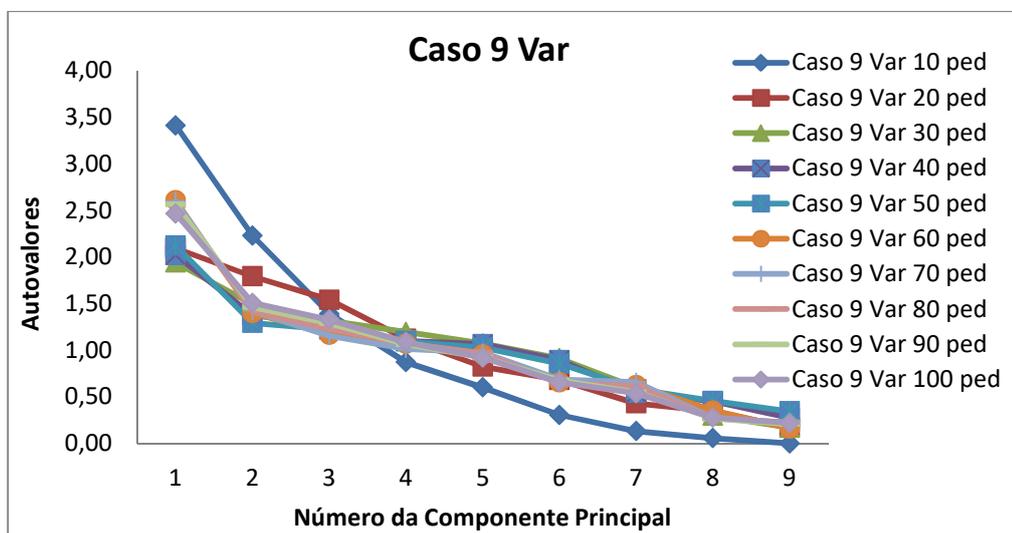


Figura 13 – Caso 9 Var: autovalores para amostras de 10 até 100 pedidos

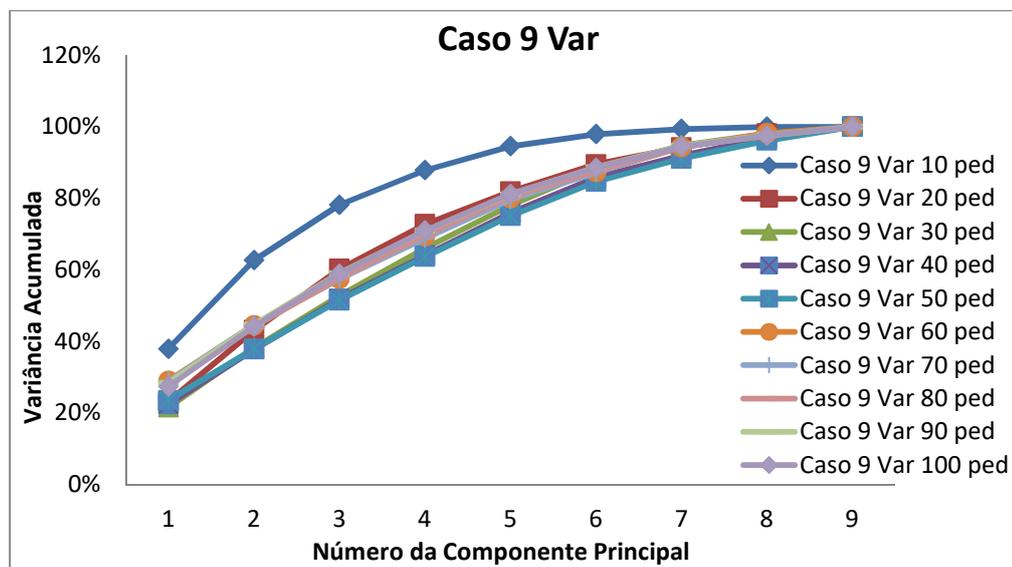


Figura 14 – Caso 9 Var: variância acumulada para amostras de 10 até 100 pedidos

Da análise da Figura 13, verifica-se que os casos com 10 e 20 pedidos se desviam dos demais, sendo que os casos com 30 até 100 pedidos são quase coincidentes, ou seja, os autovalores das componentes tendem a uma boa estabilidade a partir da amostra com 30 pedidos. Com relação às variâncias acumuladas, apenas a amostra com 10 pedidos se desvia dos demais conforme ilustrado na Figura 14.

Estes resultados evidenciam que para os casos com nove variáveis que se pretenda executar, é indicada uma amostra de pelo menos 30 pedidos de patente, preferivelmente 50 pedidos, de modo que se consiga aprovação no teste KMO e um melhor rendimento do método de ACP.

5.2.2.1.1.2 Cálculo e Avaliação das Componentes Principais

A Tabela 27 mostra os escores das quatro componentes selecionadas para cada uma das variáveis consideradas.

Tabela 27 – Escores das Componentes Principais de cada variável

Variável	Coeficiente de Ponderação (Escore)			
	Y1	Y2	Y3	Y4
Páginas de Relatório	0,4877	-0,365	0,0152	-0,0722
Páginas de Reivindicações	0,554	0,1209	-0,0843	-0,0099
Páginas de Desenhos	0,334	-0,5432	-0,1019	-0,0631
Reivindicações Independentes	0,3129	0,3875	0,2211	-0,1357
Reivindicações Dependentes	0,4362	0,413	-0,0751	0,0209
Subclassess	0,0147	0,0026	0,1338	0,9081
Ano de Depósito	-0,1427	-0,1069	0,6501	-0,3297
Inventores	0,0566	-0,4741	0,074	0,1166
Prioridades	0,178	0,0142	0,6941	0,1579

Fonte: Elaboração Própria.

Na Tabela 27 foram hachurados na cor cinza os valores considerados significativos dos fatores para cada uma das variáveis (próximos ou acima de 0,4). Em uma primeira análise, nota-se que a componente Y1 é bastante destacada das demais. Além de explicar 27,42% de toda a variância, ela está associada às principais variáveis diretamente relacionadas ao volume de dados (número de páginas de relatório, número de páginas de reivindicações e número de reivindicações dependentes). Além disso, as outras duas variáveis de volume direto (número de páginas de desenhos e número de reivindicações independentes) também apresentam valores razoáveis, acima de 0,3. Em suma, a componente Y1 é quase que uma componente totalmente associada ao volume de dados. Já a componente Y2 complementa a componente Y1 com as demais variáveis relacionadas ao volume de dados (numero de páginas de desenhos e número de reivindicações independentes), e somada à componente Y1, responde por 44,17% da variância. Tal fato se encontra em consonância com a nossa hipótese

inicial 1 das variáveis de volume. Por outro lado, as componentes Y3 e Y4 representam basicamente as influências das variáveis ano de depósito, número de prioridades e número de subclasses. São componentes que aparentam estar associadas às estratégias de desenvolvimento e gestão dos depositantes e maturidade da tecnologia envolvida, possivelmente representando influências específicas da área tecnológica dos pedidos de patente. Este resultado se encontra em consonância com a nossa hipótese inicial 2 relativa às variáveis com influência complementar ou indireta.

5.2.2.1.1.3 Correlações com Tempo

A Figura 15 mostra o gráfico das correlações do IGC e IGC_{Y1} com o tempo para o Caso 9 Var com a variação da amostra padrão de 10 em 10 pedidos de patente.

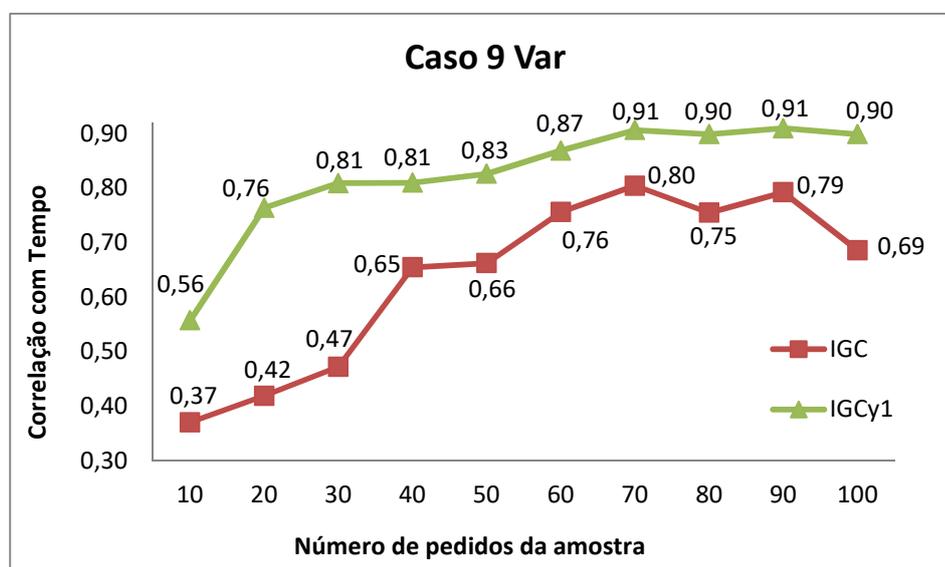


Figura 15 – Caso 9 Var: correlações do IGC com o tempo

Da análise da Figura 15, verifica-se que as correlações do IGC com o tempo crescem até a amostra de 60 pedidos, a partir da qual parece oscilar em correlações entre 0,7 e 0,8. Por outro lado, quando se analisa o IGC_{Y1} , ou seja, considerando apenas a componente principal mais significativa, relacionada basicamente às principais variáveis de volume direto, os valores também são crescentes até a amostra com 60 pedidos, sendo que a partir disso ocorre uma clara estabilidade, sendo que a correlação com o tempo atinge valores bem elevados (0,90 e 0,91). Vale ainda ressaltar que a correlação do IGC_{Y1} com o tempo já atinge altas correlações (a partir de 0,81) quando a amostra passa de 10 para 30 pedidos. Este resultado se mostra coerente com o perfil dos autovalores e variâncias acumuladas analisados nas Figuras

13 e 14, reforçando a necessidade de uma amostra com pelo menos 30 pedidos, e ainda mostra coerência com os testes KMO, com aprovação a partir de amostras com 50 pedidos.

A variável que isoladamente apresentou elevadas correlações com o tempo, com ditas correlações permanecendo estáveis ao longo de todas as simulações (oscilando entre 0,75 e 0,82), foi o número de páginas de reivindicações. A figura 15 mostra as correlações do IGC_{Y1} e da variável número de páginas de reivindicações com o tempo para o Caso 9 Var, com a variação da amostra padrão de 10 em 10 pedidos de patente.

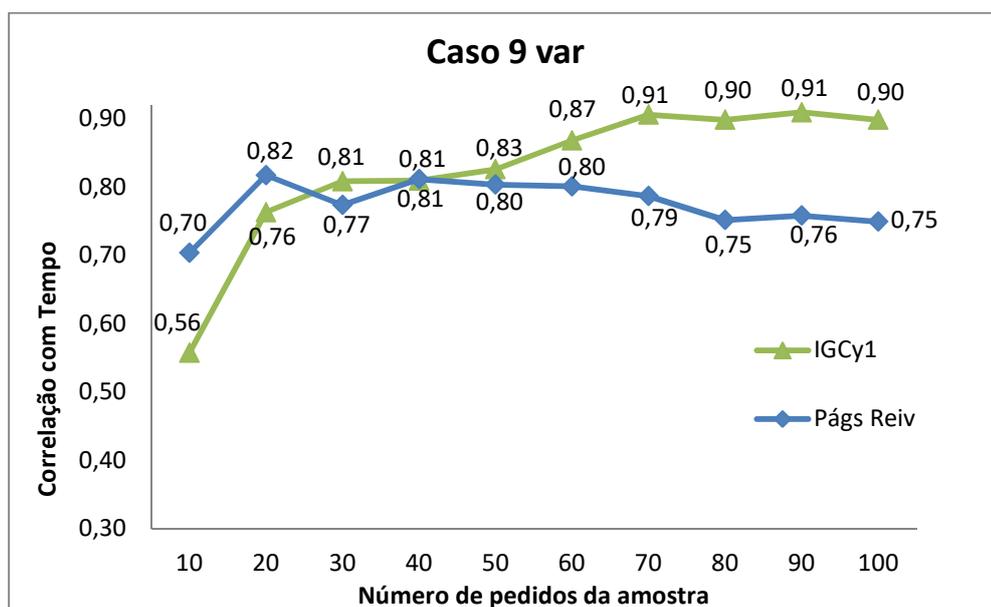


Figura 16 – Caso 9 Var: correlações do IGC_{Y1} , número de páginas de reivindicações e número total de páginas com o tempo

Da análise da Figura 15, verifica-se que a partir de uma amostra com um total de 50 pedidos de patente inicia-se uma consistente vantagem de correlação com o tempo para o IGC_{Y1} .

5.2.2.1.2 Caso 2: Caso 5 Var

O segundo caso com especial atenção de nossa parte foi o caso com as cinco principais variáveis de volume direto, denominado Caso 5 Var, incluindo as seguintes variáveis:

- ✓ número de páginas do relatório descritivo (variável 1);
- ✓ número de páginas de reivindicações (variável 2);
- ✓ número de páginas de desenhos (variável 3);
- ✓ número de reivindicações independentes (variável 4); e
- ✓ número de reivindicações dependentes (variável 5).

5.2.2.1.2.1 Cálculo da Matriz de Correlação e Seleção das Componentes Principais

Como este caso apresenta cinco variáveis, as matrizes S e R das variáveis padronizadas agora são iguais e quadradas de dimensão 5 x 5. Da mesma forma, como foram executados dez casos diferentes (com 10 até 100 pedidos), a título de ilustração é mostrada a seguir a matriz S=R relativa à amostra completa, com 100 pedidos.

$$S = R = \begin{bmatrix} 1 & 0,61 & 0,52 & 0,16 & 0,22 \\ 0,61 & 1 & 0,26 & -0,03 & -0,03 \\ 0,52 & 0,26 & 1 & 0,34 & 0,70 \\ 0,16 & -0,03 & 0,34 & 1 & 0,37 \\ 0,22 & -0,03 & 0,70 & 0,37 & 1 \end{bmatrix}$$

Tendo obtido as matrizes S e R foram então aplicadas as equações (7) a (9) para determinação dos autovalores ($\lambda_1, \lambda_2, \dots, \lambda_5$) e da matriz V de autovetores associados (escores das componentes).

De posse dos autovalores e autovetores, são então aplicadas as equações (10) e (11) para determinação das componentes principais e das suas respectivas contribuições de variâncias. É ainda calculada a contribuição acumulada das componentes (% de variância acumulada). A Tabela 28 mostra estes resultados.

Tabela 28 – Autovalores e Contribuições de Variância das Componentes Principais

Componente Principal	Autovalores	Variância	Variância Acumulada
Y1	2,34	47%	47%
Y2	1,39	28%	74%
Y3	0,71	14%	89%
Y4	0,34	7%	96%
Y5	0,22	4%	100%

Fonte: Elaboração Própria.

Da análise da Tabela 28, tanto no caso de utilização do critério considerando autovalores maiores do que um, quanto pela variância acumulada acima de 70%, seriam selecionadas as duas primeiras componentes para a determinação do IGC. Esse mesmo fenômeno foi observado para todas simulações com variações das amostras, de 10 até 100 pedidos.

Os testes de Bartlett e KMO foram aplicados para cada um dos casos. Quanto ao Teste de Bartlett, com exceção do caso com somente 10 amostras, todos os demais foram aprovados. Além disso, todas as amostras foram aprovadas no teste KMO, desde 10 até 100 pedidos. Sendo assim, verifica-se que, conforme esperado, o caso 5 Var necessita de bem

menos amostras que o caso 9 Var. As Figuras 17 e 18 mostram, respectivamente, os autovalores e as variâncias acumuladas para cada um dos casos 5 var, com amostras de 10 até 100 pedidos.

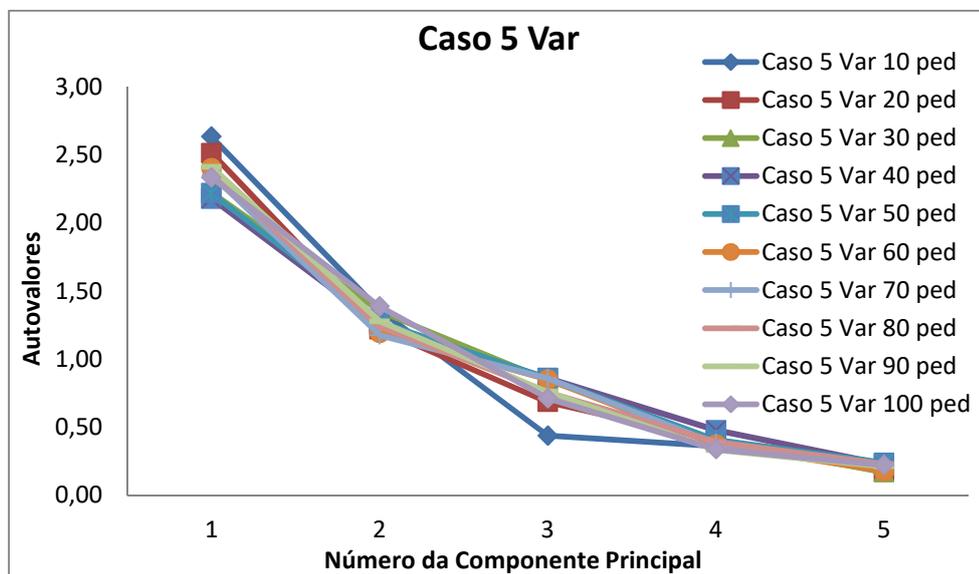


Figura 17 – Caso 5 Var: autovalores para amostras de 10 até 100 pedidos

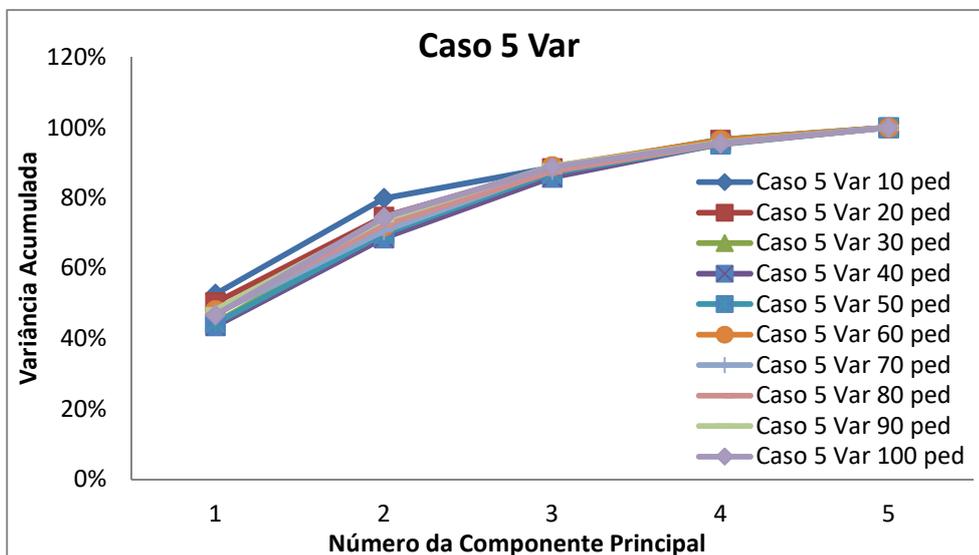


Figura 18 – Caso 5 Var: variância acumulada para amostras de 10 até 100 pedidos

Da análise da Figura 17, verifica-se que somente o caso com amostra de 10 pedidos (único não aprovado no teste de Bartlett) se desvia um pouco dos demais, sendo que o caso com 20 pedidos desvia ligeiramente e os casos com 30 até 100 pedidos são quase coincidentes, ou seja, os autovalores das componentes já tendem a uma boa estabilidade a

partir da amostra com 20 pedidos. O mesmo fenômeno pode ser observado ao analisar as variâncias acumuladas das amostras conforme ilustrado na Figura 18.

Estes resultados evidenciam que para os casos com cinco variáveis que se pretenda executar, é indicada uma amostra de pelo menos 20 pedidos de patente, preferivelmente 30 pedidos, de modo que se consiga um melhor rendimento do método de ACP.

5.2.2.1.2.2 Cálculo e Avaliação das Componentes Principais

A Tabela 29 mostra os escores das duas primeiras componentes principais para cada uma das variáveis consideradas do caso 5 var com 100 pedidos.

Tabela 29 – Escores das Componentes Principais de cada variável

Variáveis	Coeficiente de Ponderação (Escores)	
	Y1	Y2
Páginas de Relatório	0,4868	-0,4332
Páginas de Desenhos	0,3170	-0,6567
Páginas de Reivindicações	0,5798	0,1349
Reivindicações Independentes	0,3282	0,4107
Reivindicações Dependentes	0,4676	0,4407

Fonte: Elaboração Própria.

Na Tabela 29 foram hachurados na cor cinza os valores considerados singificativos dos fatores para cada uma das variáveis (próximos ou acima de 0,4). Em uma primeira análise, nota-se que a componente Y1 é bastante destacada das demais. Além de explicar 47% de toda a variância, ela está associada às principais variáveis diretamente relacionadas ao volume de dados (número de páginas de relatório, número de páginas de reivindicações e número de reivindicações dependentes). Além disso, as outras duas variáveis de volume direto (número de páginas de desenhos e número de reivindicações independentes) também apresentam valores positivos e razoáveis, acima de 0,3. Em suma, a componente Y1 é quase que uma componente totalmente associada ao volume de dados. Já a componente Y2 complementa a componente Y1 com valores elevados de escores das demais variáveis relacionadas ao volume de dados (páginas de desenhos e número de reivindicações independentes) e, somada à componente Y1, responde por 74% da variância. Tal fato também se encontra em consonância com a nossa hipótese inicial 1.

5.2.2.1.2.3 Correlações com o Tempo

A Figura 19 mostra o gráfico das correlações do IGC e IGC_{Y1} com o tempo para o Caso 5 Var com a variação da amostra padrão de 10 em 10 pedidos de patente. Além disso,

para efeito de comparação, é incluída a correlação do número de páginas de reivindicações com o tempo.

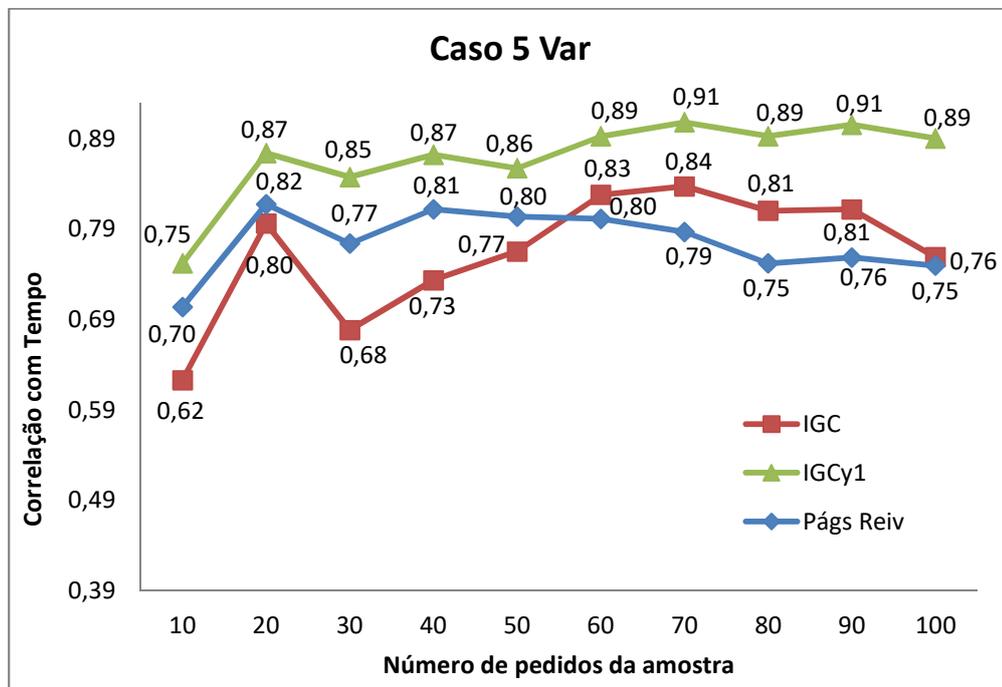


Figura 19 – Caso 5 Var: correlações do IGC_{Y1} , IGC, número de páginas de reivindicações e número total de páginas com o tempo

Da análise da Figura 19 verifica-se que, a partir da amostra com 20 pedidos, a curva de correlação do IGC_{Y1} com tempo apresenta boa estabilidade e valores bem elevados (cerca de 0,89). Verifica-se ainda um perfil bem semelhante ao da curva de número de páginas de reivindicações, que também se estabiliza a partir da amostra de 20 pedidos, mas com correlações mais baixas (entre 0,75 e 0,80). Já a curva de correlação do IGC com o tempo apresenta uma oscilação maior, com tendência de estabilidade apenas a partir da amostra com 60 pedidos e também com correlações mais baixas que a do IGC_{Y1} . Vale ressaltar ainda que a correlação do IGC_{Y1} e do número de páginas de reivindicações com o tempo aumenta bastante quando a amostra passa de 10 para 20 pedidos (respectivamente de 0,87 e 0,82). Tais resultados se mostram coerentes com os perfis dos autovalores e variâncias acumuladas analisados nas Figuras 17 e 18, reforçando a necessidade de uma amostra com pelo menos 20 pedidos. Por fim, mais uma vez os resultados sugerem que o IGC_{Y1} apresenta as maiores correlações com o tempo.

5.2.2.1.3 Caso 3: Caso 4 Var

O terceiro caso em análise foi montado a partir do Caso 5 Var, excluindo a variável número de páginas de desenhos, a que apresentou menor correlação com tempo entre as ditas variáveis de volume direto. Consequentemente, este caso foi denominado de Caso 4 Var, incluindo as seguintes variáveis:

- ✓ número de páginas do relatório descritivo (variável 1);
- ✓ número de páginas de reivindicações (variável 2);
- ✓ número de reivindicações independentes (variável 3); e
- ✓ número de reivindicações dependentes (variável 4).

5.2.2.1.3.1 Cálculo da Matriz de Correlação e Seleção das Componentes Principais

Como este caso apresenta quatro variáveis, as matrizes S e R das variáveis padronizadas agora são iguais e quadradas de dimensão 4 x 4. Da mesma forma, como foram executados dez casos diferentes (com 10 até 100 pedidos), a título de ilustração é mostrada a seguir a matriz S=R relativa à amostra completa, com 100 pedidos.

$$S = R = \begin{bmatrix} 1 & 0,52 & 0,16 & 0,22 \\ 0,52 & 1 & 0,34 & 0,70 \\ 0,16 & 0,34 & 1 & 0,37 \\ 0,22 & 0,70 & 0,37 & 1 \end{bmatrix}$$

Tendo obtido as matrizes S e R foram então aplicadas as equações (7) a (9) para determinação dos autovalores ($\lambda_1, \lambda_2, \lambda_3$ e λ_4) e da matriz V de autovetores associados (escores das componentes).

De posse dos autovalores e autovetores, são então aplicadas as equações (10) e (11) para determinação das componentes principais e das suas respectivas contribuições de variâncias. É ainda calculada a contribuição acumulada das componentes (% de variância acumulada). A Tabela 30 mostra estes resultados.

Tabela 30 – Autovalores e Contribuições de Variância das Componentes Principais

Componente Principal	Autovalores	Variância	Variância Acumulada
Y1	2,21	55%	55%
Y2	0,89	22%	77%
Y3	0,68	17%	94%
Y4	0,22	6%	100%

Fonte: Elaboração Própria.

Da análise da Tabela 30, verifica-se que no caso de utilização do critério considerando autovalores maiores do que um, apenas a componente Y1 seria selecionada, ou seja, o $IGC_{\lambda>1}$ seria igual ao IGC_{Y1} . Utilizando o critério de 70% da variância acumulada, seriam selecionadas as duas primeiras componentes para a determinação do $IGC_{70\%}$. Esse mesmo fenômeno foi observado para todas simulações com variações das amostras, de 10 até 100 pedidos.

Os testes de Bartlett e KMO foram aplicados para cada um dos casos. Quanto ao Teste de Bartlett, com exceção do caso com somente 10 amostras, todos os demais foram aprovados. Além disso, todas as amostras foram aprovadas no teste KMO, desde 10 até 100 pedidos. Sendo assim, verificou-se resultados similares aos do Caso 5 Var.

As Figuras 20 e 21 mostram, respectivamente, os autovalores e as variâncias acumuladas para cada um dos casos 4 var, com amostras de 10 até 100 pedidos.

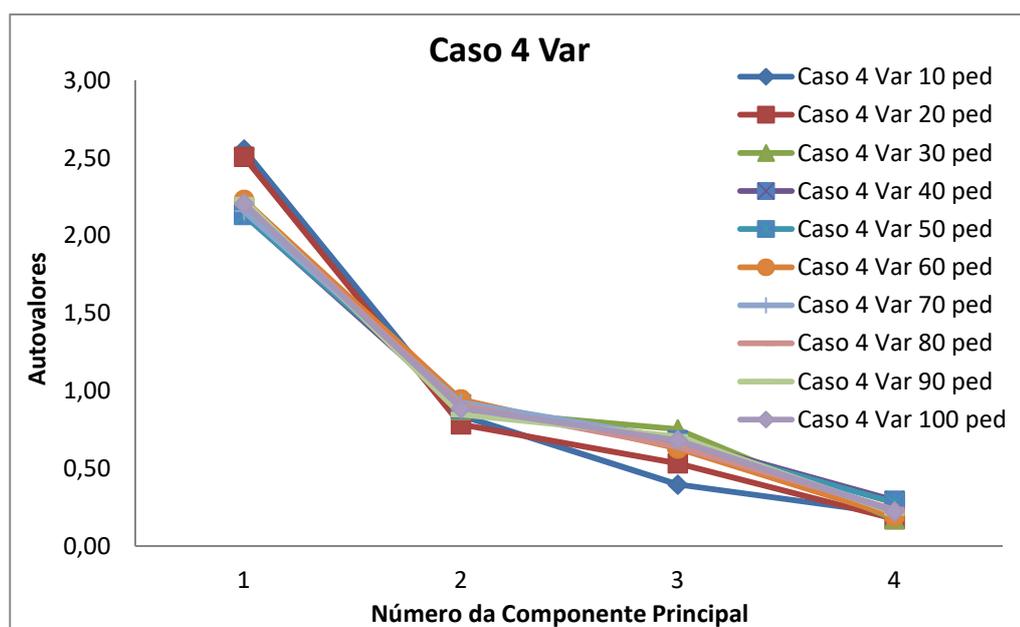


Figura 20 – Caso 4 Var: autovalores para amostras de 10 até 100 pedidos

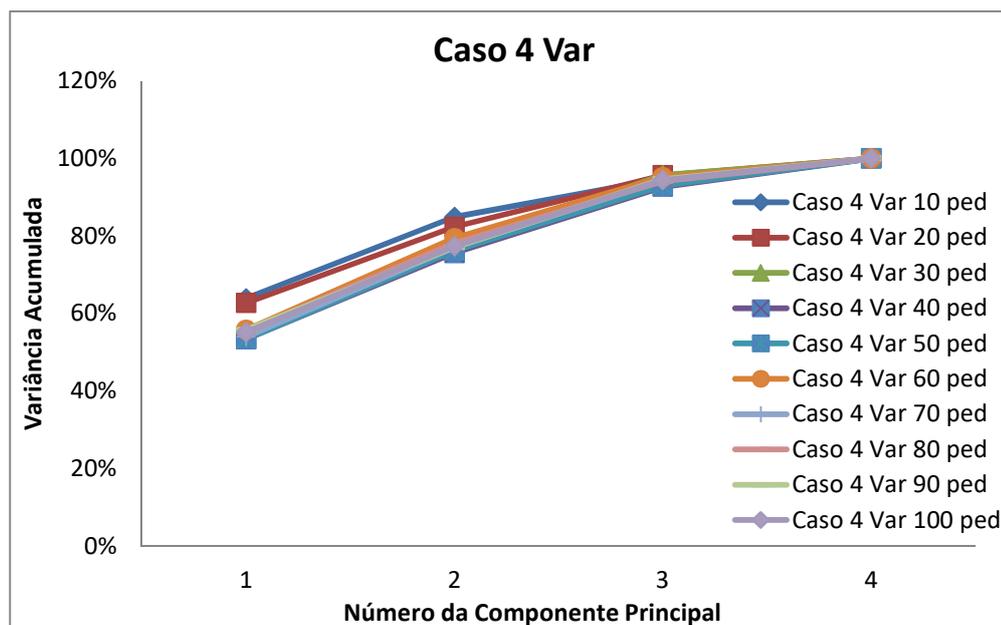


Figura 21 – Caso 4 Var: variância acumulada para amostras de 10 até 100 pedidos

Da análise da Figura 20, verifica-se que somente os casos com amostras de 10 e 20 pedidos se desviam um pouco dos demais, sendo que os casos com 30 até 100 pedidos são praticamente coincidentes, ou seja, os autovalores das componentes já tendem a uma boa estabilidade a partir da amostra com 30 pedidos. O mesmo fenômeno pode ser observado ao analisar as variâncias acumuladas das amostras conforme ilustrado na Figura 21.

Por outro lado, cabe destacar novamente que todos os casos (10 a 100 amostras) passaram nos testes de Bartlett e KMO e que, mesmo com o perfil um pouco distinto das amostras com 10 e 20 pedidos (vide figuras 20 e 21), as variâncias e autovalores apresentaram valores até melhores do que os demais. Conseqüentemente, especificamente para este caso 4 var, não deve ser necessariamente descartada a possibilidade de utilização de amostras com menos de 30 pedidos.

Portanto, os resultados evidenciam que, para os casos com quatro variáveis que se pretenda executar, é indicada uma amostra de pelo menos 10 pedidos de patente, preferivelmente 30 pedidos, de modo que se consiga um bom rendimento do método de ACP.

5.2.2.1.3.2 Cálculo e Avaliação das Componentes Principais

A Tabela 31 mostra os escores das duas primeiras componentes principais para cada uma das variáveis consideradas do caso 4 var com 100 pedidos.

Tabela 31 – Escores das Componentes Principais de cada variável

Variáveis	Coeficiente de Ponderação (Escores)	
	Y1	Y2
Páginas de Relatório	0,4149	-0,7221
Páginas de Reivindicações	0,6084	-0,1449
Reivindicações Independentes	0,3946	0,6263
Reivindicações Dependentes	0,5496	0,2556

Fonte: Elaboração Própria.

Na Tabela 31 foram hachurados na cor cinza os valores considerados singificativos dos fatores para cada uma das variáveis (próximos ou acima de 0,4). Em uma primeira análise, nota-se que a componente Y1 é bastante destacada das demais. Além de explicar 55% de toda a variância, ela está associada positivamente a quase todas as principais variáveis diretamente relacionadas ao volume de dados (número de páginas de relatório, número de páginas de reivindicações, número de reivindicações independentes e dependentes). Já a componente Y2 complementa a componente Y1 com valor bem alto de escore da variável número de reivindicações independentes e, somada à componente Y1, responde por 77% da variância.

5.2.2.1.3.3 Correlações com Tempo

A Figura 22 mostra o gráfico das correlações do IGC_{Y1} com o tempo para o Caso 4 Var com a variação da amostra padrão de 10 em 10 pedidos de patente. Além disso, para efeito de comparação, é incluída a curva de correlação do número de páginas de reivindicações.

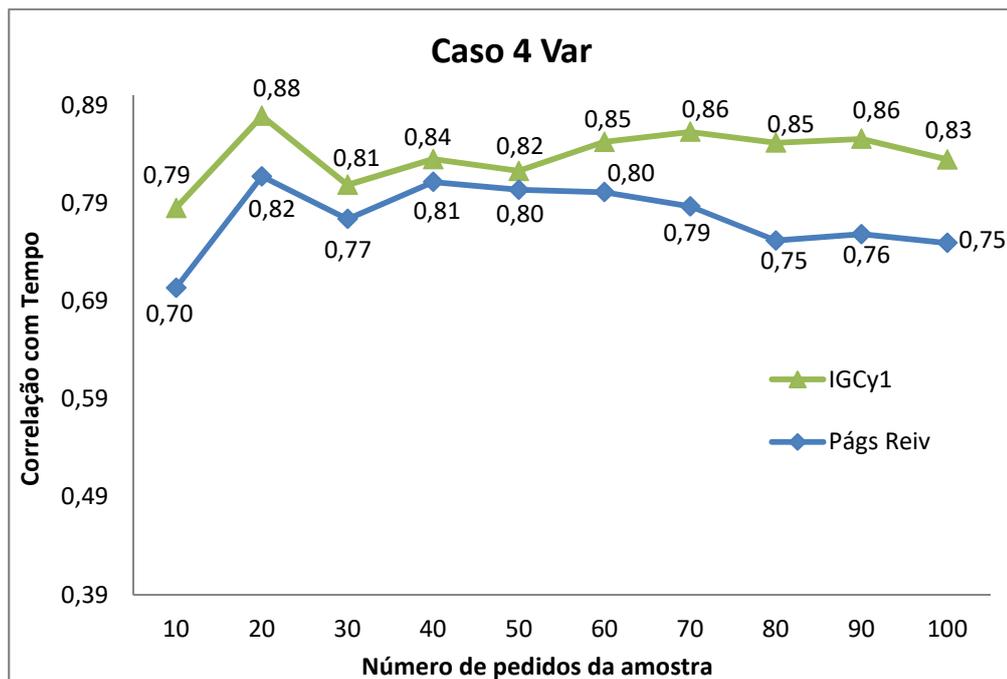


Figura 22 – Caso 4 Var: correlações do IGC_{Y1} , número de páginas de reivindicações e número total de páginas com o tempo

Da análise da Figura 22, verifica-se novamente que os perfis das curvas de correlação do IGC_{Y1} e número de páginas de reivindicações com tempo são similares. Ambas aumentam bastante a correlação com tempo ao passar de amostras com 10 para 20 pedidos, e já apresentam valores altos e mais estáveis de correlação a partir de então. Vale ressaltar também que as duas curvas apresentam uma boa correlação com tempo mesmo a partir de uma amostra com apenas 10 pedidos, resultado que se mostra coerente com o perfil dos autovalores e variâncias acumuladas analisados nas figuras 20 e 21.

Por fim, mais uma vez os resultados sugerem que o IGC_{Y1} (neste caso também igual ao $IGC_{\lambda>1}$) apresenta as maiores correlações com o tempo.

5.2.2.1.4 Caso 4: Caso 3 Var

O quarto caso em destaque foi montado a partir do Caso 4 Var, mas agregando o número de reivindicações independentes e dependentes em uma única variável, pois além de possibilitar uma comparação da influência dessa agregação, seria muito mais fácil e prático para uma aplicação real coletar o número total de reivindicações. Conseqüentemente, este caso foi denominado de Caso 3 Var, incluindo as seguintes variáveis:

- ✓ número de páginas do relatório descritivo (variável 1);

- ✓ número de páginas de reivindicações (variável 2); e
- ✓ número total de reivindicações (variável 3).

5.2.2.1.4.1 Cálculo da Matriz de Correlação e Seleção das Componentes Principais

Como este caso apresenta três variáveis, as matrizes S e R das variáveis padronizadas agora são iguais e quadradas de dimensão 3 x 3. Da mesma forma, como foram executados dez casos diferentes (com 10 até 100 pedidos), a título de ilustração é mostrada a seguir a matriz S=R relativa à amostra completa, com 100 pedidos.

$$S = R = \begin{bmatrix} 1 & 0,23 & 0,52 \\ 0,23 & 1 & 0,71 \\ 0,52 & 0,71 & 1 \end{bmatrix}$$

Tendo obtido as matrizes S e R foram então aplicadas as equações (7) a (9) para determinação dos autovalores (λ_1, λ_2 e λ_3) e da matriz V de autovetores associados (escores das componentes).

De posse dos autovalores e autovetores, são então aplicadas as equações (10) e (11) para determinação das componentes principais e das suas respectivas contribuições de variâncias. É ainda calculada a contribuição acumulada das componentes (% de variância acumulada). A Tabela 32 mostra estes resultados.

Tabela 32 – Autovalores e Contribuições de Variância das Componentes Principais

Componente Principal	Autovalores	Variância	Variância Acumulada
Y1	2,00	67%	67%
Y2	0,78	26%	93%
Y3	0,22	7%	100%

Fonte: Elaboração Própria.

Da análise da Tabela 32, verifica-se que, tanto no caso de utilização do critério considerando autovalores maiores do que um, quanto utilizando o critério de 70% (ou ao menos acima de 60%) da variância acumulada, seria selecionada somente a primeira componente para a determinação do IGC. Sendo assim, neste caso teria-se $IGC = IGC_{70\%} = IGC_{\lambda > 1} = IGC_{Y1}$. Esse mesmo fenômeno foi observado para todas simulações do caso 3 var com as amostras de 10 até 100 pedidos, com variações muito pequenas na variância da componente Y1, de 67 a 73%. Estes resultados evidenciam um ótimo rendimento do método

de ACP, apresentando uma excelente estabilidade das componentes e reduzindo sempre o problema para apenas uma dimensão.

Os testes de Bartlett e KMO foram aplicados para cada um dos casos. Quanto ao Teste de Bartlett, com exceção do caso com somente 10 amostras, todos os demais foram aprovados. Além disso, todas as amostras foram aprovadas no teste KMO, desde 10 até 100 pedidos.

As Figuras 23 e 24 mostram, respectivamente, os autovalores e as variâncias acumuladas para cada um dos casos 3 var, com amostras de 10 até 100 pedidos.

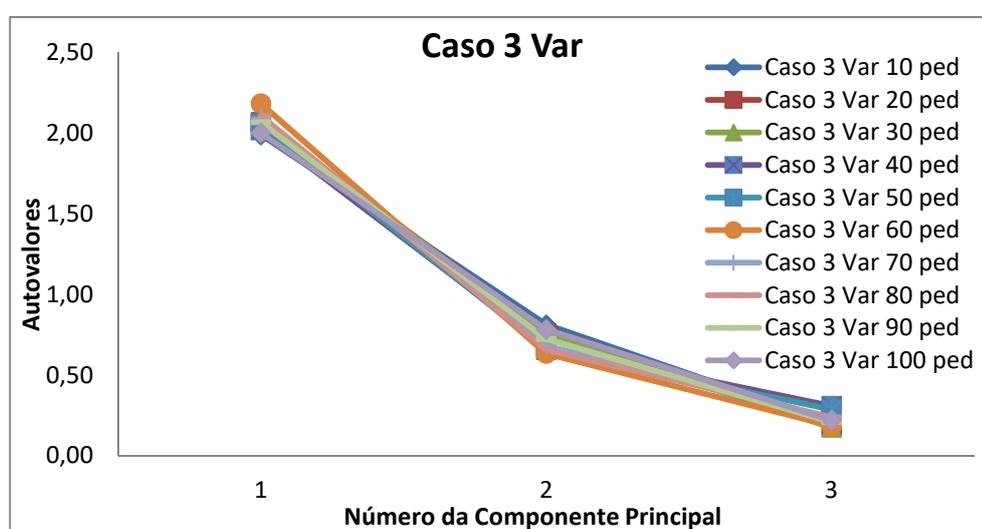


Figura 23 – Caso 3 Var: autovalores para amostras de 10 até 100 pedidos

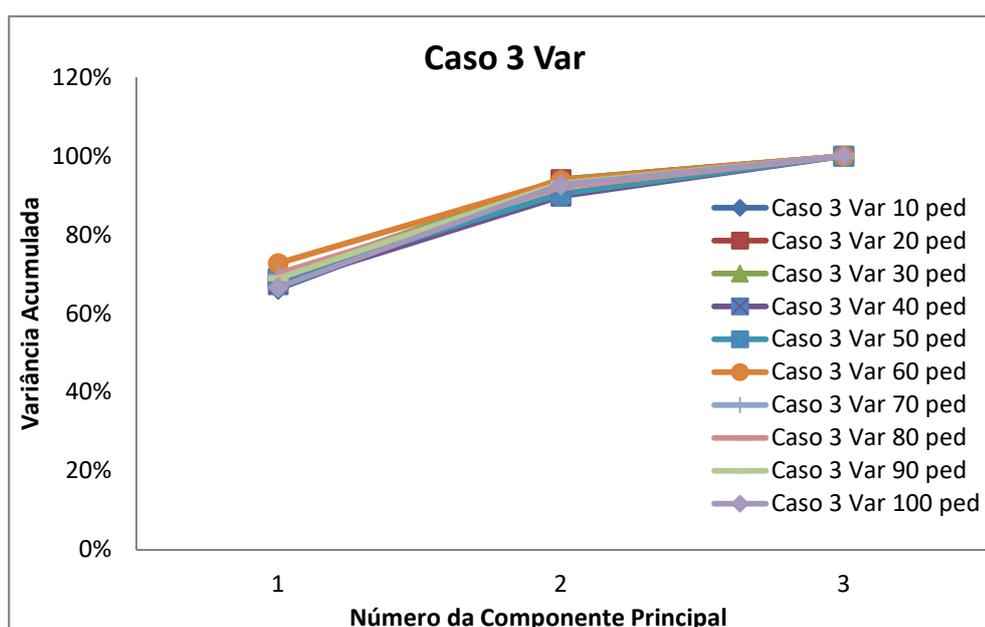


Figura 24 – Caso 3 Var: variância acumulada para amostras de 10 até 100 pedidos

Da análise da Figura 23, verifica-se que todos os casos, desde amostras de 10 até 100 pedidos, são praticamente coincidentes, ou seja, os autovalores das componentes apresentam uma excelente estabilidade. O mesmo fenômeno pode ser observado ao analisar as variâncias acumuladas das amostras conforme ilustrado na Figura 24.

Estes resultados evidenciam que, para os casos com três variáveis que se pretenda executar, já é possível utilizar uma amostra de 10 pedidos de patente, sendo preferível a partir de 20 pedidos, de modo que se consiga um melhor rendimento do método de ACP.

5.2.2.1.4.2 Cálculo e Avaliação das Componentes Principais

A Tabela 33 mostra os escores das duas primeiras componentes principais para cada uma das variáveis consideradas do caso 3 var com 100 pedidos.

Tabela 33 – Escores das Componentes Principais de cada variável

Variáveis	Coeficiente de Ponderação (Escore)	
	Y1	Y2
Páginas de Relatório	0,4776	-0,8164
Total de Reivindicações	0,5793	0,5705
Páginas de Reivindicações	0,6606	0,0895

Fonte: Elaboração Própria.

Na Tabela 33 foram hachurados na cor cinza os valores considerados significativos dos fatores para cada uma das variáveis (próximos ou acima de 0,4). Em uma primeira análise, nota-se que a componente Y1 é bastante destacada das demais, sendo responsável por explicar 67% de toda a variância. Neste caso, ambas as componentes Y1 e Y2 estão associadas às principais variáveis diretamente relacionadas ao volume de dados, sendo que Y1 está positivamente associada com as três variáveis em análise (número de páginas de relatório, número de páginas de reivindicações e número total de reivindicações). Este resultado mais uma vez evidencia que as variáveis de volume direto, basicamente representadas pela componente principal Y1, são responsáveis pela maior parte da variância dos dados dos pedidos de patente, tendendo a representar de maneira satisfatória sua complexidade.

5.2.2.1.4.3 Correlações com o Tempo

A Figura 25 mostra o gráfico das correlações do IGC com o tempo para o Caso 3 Var com a variação da amostra padrão de 10 em 10 pedidos de patente. Além disso, para efeito de comparação, é incluída a correlação do número de páginas de reivindicações com o tempo.

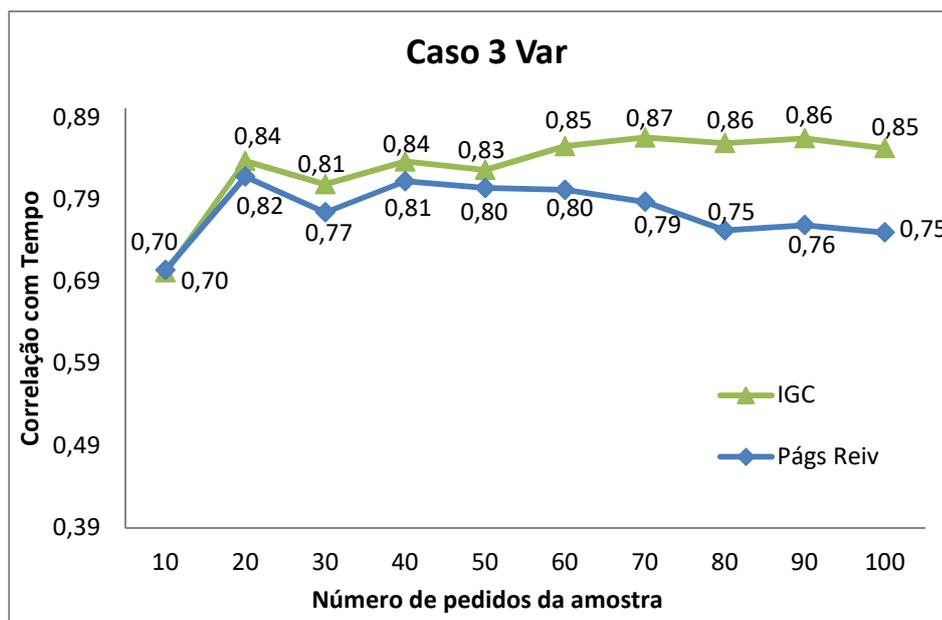


Figura 25 – Caso 3 Var: correlações do IGC com o tempo

Da análise da Figura 25, verifica-se novamente que os perfis das curvas de correlação do IGC e número de páginas de reivindicações com tempo são similares. Ambas aumentam bastante a correlação com tempo ao passar de amostras com 10 para 20 pedidos, e já apresentam valores altos e mais estáveis de correlação a partir de então. Vale ressaltar também que as duas curvas apresentam uma boa correlação com tempo mesmo a partir de uma amostra com apenas 10 pedidos, resultado que se mostra coerente com o perfil dos autovalores e variâncias acumuladas analisados nas figuras 23 e 24.

Por fim, mais uma vez os resultados sugerem que o IGC (neste caso $IGC = IGC_{70\%} = IGC_{\lambda > 1} = IGC_{Y1}$) apresenta as maiores correlações com o tempo.

5.2.2.1.5 Caso 5: Caso 3 Var (2)

Por fim, foi selecionado um último e quinto caso particular de análise, também com três variáveis, mas substituindo o número de páginas do relatório descritivo pelo número total de páginas do pedido (soma das páginas do relatório, reivindicações e desenhos), dado que esta última teve uma maior correlação individual com o tempo. Portanto, este caso foi denominado Caso 3 Var (2), incluindo as seguintes variáveis:

- ✓ número total de páginas do pedido (variável 1);
- ✓ número de páginas de reivindicações (variável 2); e
- ✓ número total de reivindicações (variável 3).

5.2.2.1.5.1 Cálculo da Matriz de Correlação e Seleção das Componentes Principais

Como este caso apresenta três variáveis, as matrizes S e R das variáveis padronizadas agora são iguais e quadradas de dimensão 3 x 3. Da mesma forma, como foram executados dez casos diferentes (com 10 até 100 pedidos), a título de ilustração é mostrada a seguir a matriz S=R relativa à amostra completa, com 100 pedidos.

$$S = R = \begin{bmatrix} 1 & 0,24 & 0,57 \\ 0,24 & 1 & 0,71 \\ 0,57 & 0,71 & 1 \end{bmatrix}$$

Tendo obtido as matrizes S e R foram então aplicadas as equações (7) a (9) para determinação dos autovalores (λ_1, λ_2 e λ_3) e da matriz V de autovetores associados (escores das componentes).

De posse dos autovalores e autovetores, são então aplicadas as equações (10) e (11) para determinação das componentes principais e das suas respectivas contribuições de variâncias. É ainda calculada a contribuição acumulada das componentes (% de variância acumulada). A Tabela 34 mostra estes resultados.

Tabela 34 – Autovalores e Contribuições de Variância das Componentes Principais

Componente Principal	Autovalores	Variância	Variância Acumulada
Y1	2,04	68%	68%
Y2	0,77	26%	94%
Y3	0,20	7%	100%

Fonte: Elaboração Própria.

Da análise da Tabela 34, verifica-se que, tanto no caso de utilização do critério considerando autovalores maiores do que um, quanto utilizando o critério de 70% (ao menos acima de 60%) da variância acumulada, seria selecionada somente a primeira componente para a determinação do IGC. Sendo assim, neste caso teríamos $IGC = IGC_{70\%} = IGC_{\lambda > 1} = IGC_{Y1}$. Esse mesmo fenômeno foi observado para todas as simulações do caso 3 var (2) com as amostras de 10 até 100 pedidos, com variações muito pequenas na variância da componente Y1, de 68 a 72%. Assim como no caso 3 var anterior, estes resultados evidenciam um ótimo rendimento do método de ACP, apresentando uma excelente estabilidade das componentes e reduzindo sempre o problema para apenas uma dimensão.

Os testes de Bartlett e KMO foram aplicados para cada um dos casos. Quanto ao Teste de Bartlett, com exceção do caso com somente 10 amostras, todos os demais foram aprovados. Além disso, todas as amostras foram aprovadas no teste KMO, desde 10 até 100 pedidos.

As Figuras 26 e 27 mostram, respectivamente, os autovalores e as variâncias acumuladas para cada um dos casos 3 var (2), com amostras de 10 até 100 pedidos.

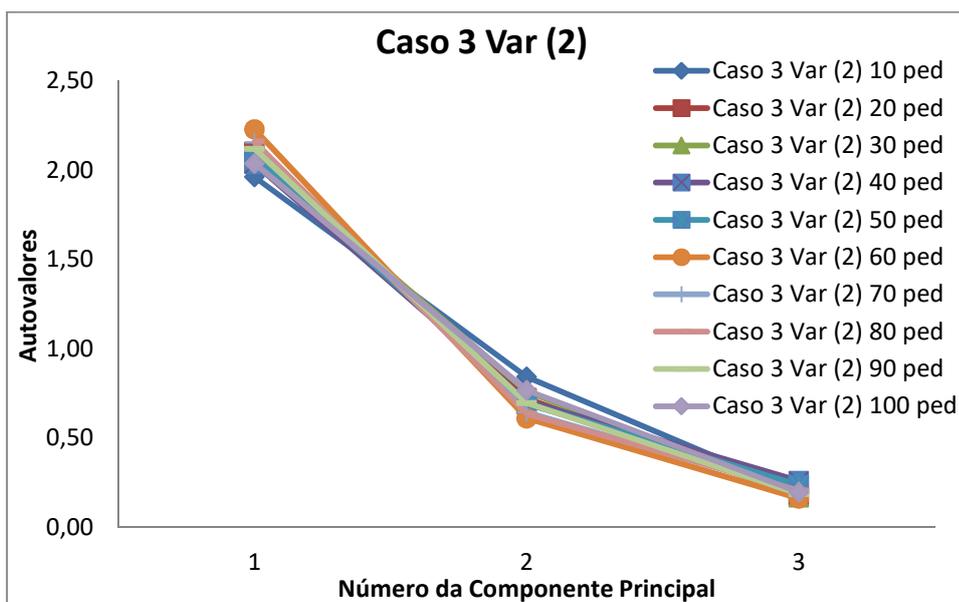


Figura 26 – Caso 3 Var (2): autovalores para amostras de 10 até 100 pedidos

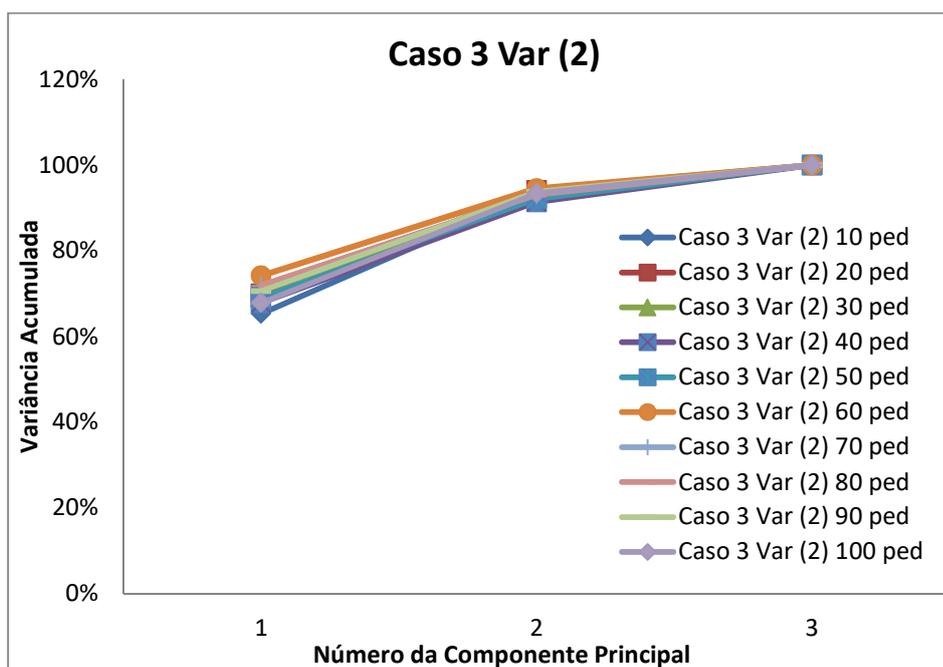


Figura 27 – Caso 3 Var (2): variância acumulada para amostras de 10 até 100 pedidos

Da análise da Figura 26, verifica-se que todos os casos, desde amostras de 10 até 100 pedidos, são praticamente coincidentes, ou seja, os autovalores das componentes apresentam uma excelente estabilidade. O mesmo fenômeno pode ser observado ao analisar as variâncias acumuladas das amostras conforme ilustrado na Figura 27.

Estes resultados reforçam que, para os casos com três variáveis que se pretenda executar, já é possível utilizar uma amostra de 10 pedidos de patente, sendo preferível a partir de 20 pedidos, de modo que se consiga um melhor rendimento do método de ACP.

5.2.2.1.5.2 Cálculo e Avaliação das Componentes Principais

A Tabela 35 mostra os escores das duas primeiras componentes principais para cada uma das variáveis consideradas do caso 3 var (2) com 100 pedidos.

Tabela 35 – Escores das Componentes Principais de cada variável

Variáveis	Coeficiente de Ponderação (Escore)	
	Y1	Y2
Total de Páginas	0,4940	-0,7885
Total de Reivindicações	0,5665	0,6115
Páginas de Reivindicações	0,6595	0,0655

Fonte: Elaboração Própria.

Na Tabela 35 foram hachurados na cor cinza os valores considerados significativos dos fatores para cada uma das variáveis (próximos ou acima de 0,4). Em uma primeira análise, nota-se que a componente Y1 destaca-se das demais, sendo responsável por explicar 68% de toda a variância. Neste caso, ambas as componentes Y1 e Y2 estão associadas às principais variáveis diretamente relacionadas ao volume de dados, sendo que Y1 está positivamente associada às três variáveis em análise (número de total de páginas, número de páginas de reivindicações e número total de reivindicações). Vale ressaltar ainda que este resultado é praticamente idêntico ao do caso 3 var e, mais uma vez, evidencia que as variáveis de volume direto, basicamente representadas pela componente principal Y1, são responsáveis pela maior parte da variância dos dados dos pedidos de patente, tendendo a representar de maneira satisfatória sua complexidade.

5.2.2.1.5.3 Correlações com Tempo

A Figura 28 mostra o gráfico das correlações do IGC com o tempo para o Caso 3 Var (2) com a variação da amostra padrão de 10 em 10 pedidos de patente. Além disso, para efeito de comparação, é incluída a correlação do número de páginas de reivindicações com o tempo.

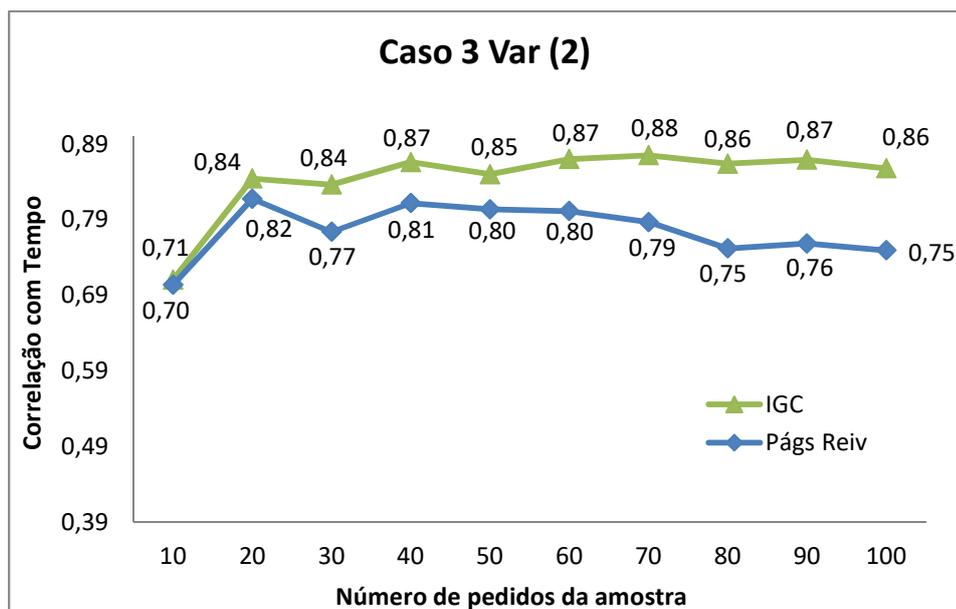


Figura 28 – Caso 3 Var (2): correlações do IGC com o tempo

Da análise da Figura 28, verifica-se novamente que os perfis das curvas de correlação do IGC e número de páginas de reivindicações com tempo são similares. Ambas aumentam bastante a correlação com tempo ao passar de amostras com 10 para 20 pedidos, e já apresentam valores altos e mais estáveis de correlação a partir de então. Vale ressaltar também que as duas curvas apresentam uma boa correlação com tempo mesmo a partir de uma amostra com apenas 10 pedidos, resultado que se mostra coerente com o perfil dos autovalores e variâncias acumuladas analisados nas figuras 26 e 27. Destaque-se ainda que, a partir da amostra com 20 pedidos, a correlação do IGC com o tempo fica sempre acima de 0,84, chegando valores de até 0,87.

Por fim, cabe destacar que este resultado é praticamente idêntico ao dos casos 3 var e 4 var e, mais uma vez, os resultados sugerem que o IGC (neste caso $IGC = IGC_{70\%} = IGC_{\lambda>1} = IGC_{Y1}$) apresenta as maiores correlações com o tempo, levando vantagem com relação a todas as variáveis individualmente. Além de apresentar uma alta correlação com o tempo, este é o caso com maior indicação para aplicação prática, pois, além de apresentar somente três variáveis e componentes principais com alta estabilidade, não necessita da separação das reivindicações em independentes e dependentes, facilitando bastante a coleta de dados.

5.2.2.2 Análise de Sensibilidade do IGC: comparação dos casos

Após análise do comportamento e da convergência do IGC dentro de cada um dos casos, adicionalmente, cabe analisar e comparar os IGC e IGC_{Y1} para todos os casos

estudados entre si, de modo a verificar explicitamente o comportamento quando da variação do número de variáveis (9, 5, 4 e 3) e do tamanho das amostras (10 a 100 pedidos de patente).

As Figuras 29 a 38 apresentam os gráficos das correlações do IGC e IGC_{Y1} com tempo para cada uma das amostras estudadas (10, 20, 30, ... , 90 e 100), sendo que no eixo x horizontal estão marcados os números de variáveis de cada um dos casos (9, 5, 4, 3 e 3). Vale ressaltar que, como há dois casos com três variáveis, há dois pontos com o valor três (3) no eixo x, sendo o penúltimo relativo ao Caso 3 Var e o último relativo ao Caso 3 Var (2).

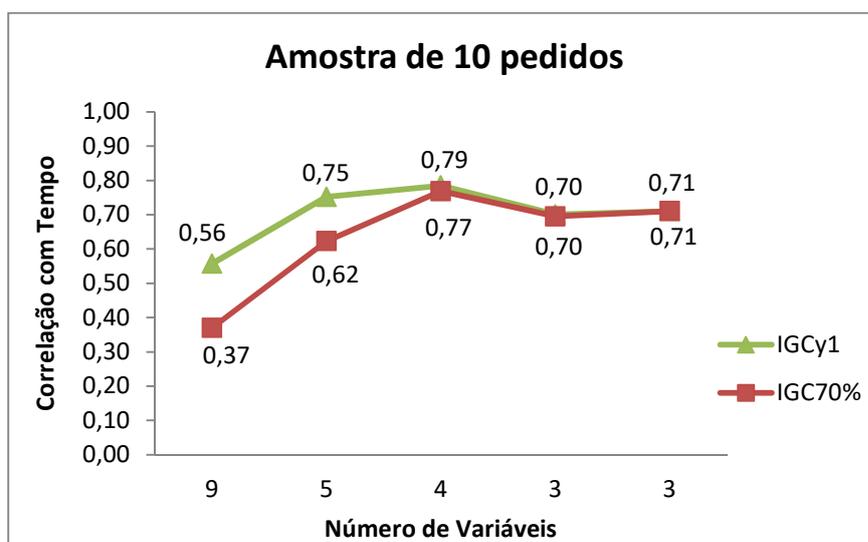


Figura 29 – Correlações do IGC e IGC_{Y1} com o tempo para amostras de 10 pedidos

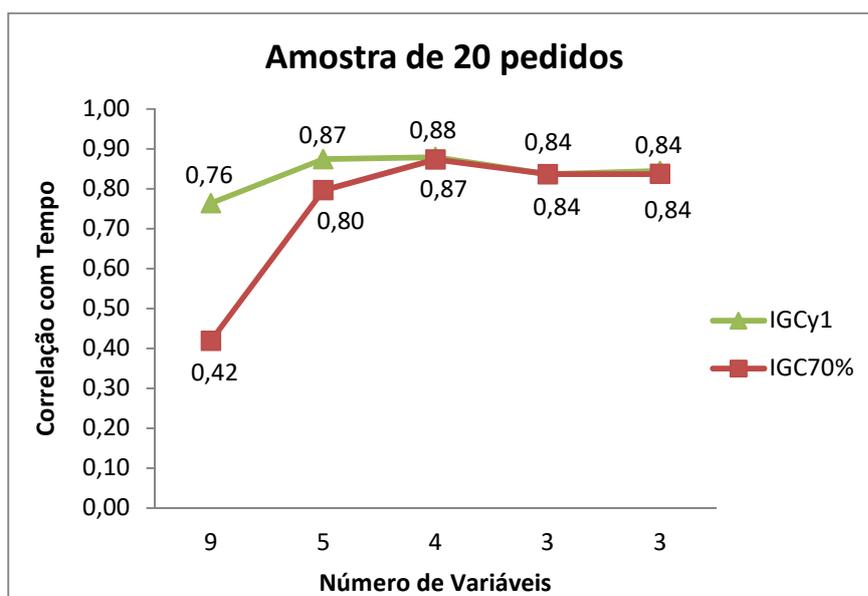


Figura 30 – Correlações do IGC e IGC_{Y1} com o tempo para amostras de 20 pedidos

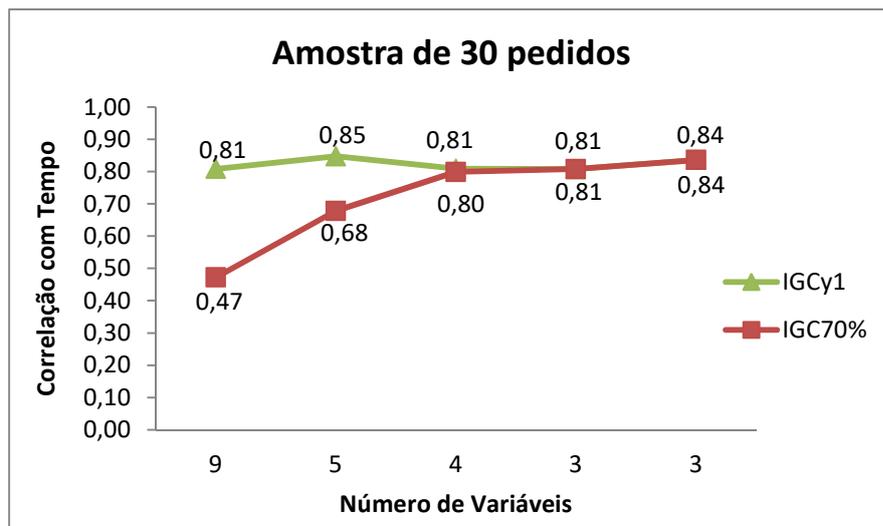


Figura 31 – Correlações do IGC e IGC_{Y1} com o tempo para amostras de 30 pedidos

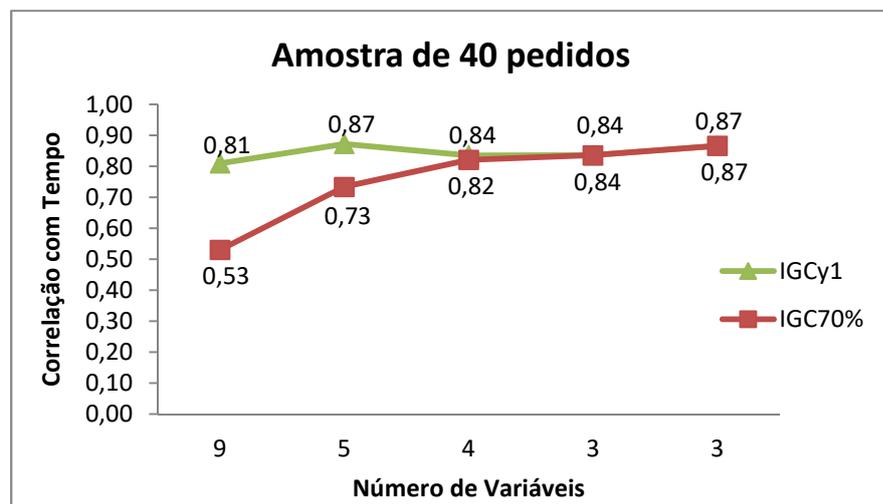


Figura 32 – Correlações do IGC e IGC_{Y1} com o tempo para amostras de 40 pedidos

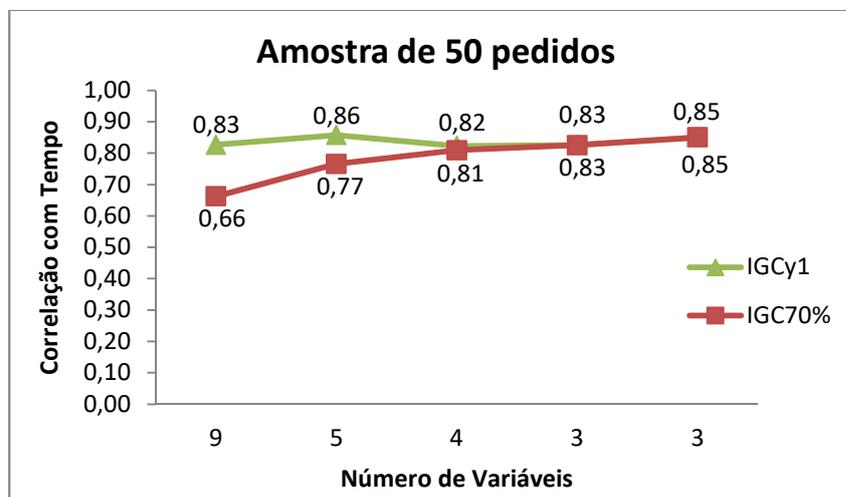


Figura 33 – Correlações do IGC e IGC_{Y1} com o tempo para amostras de 50 pedidos

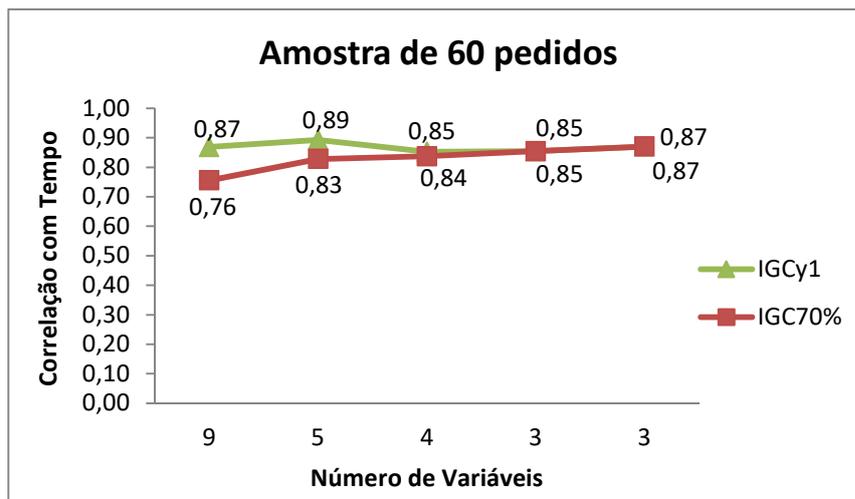


Figura 34 – Correlações do IGC e IGC_{Y1} com o tempo para amostras de 60 pedidos

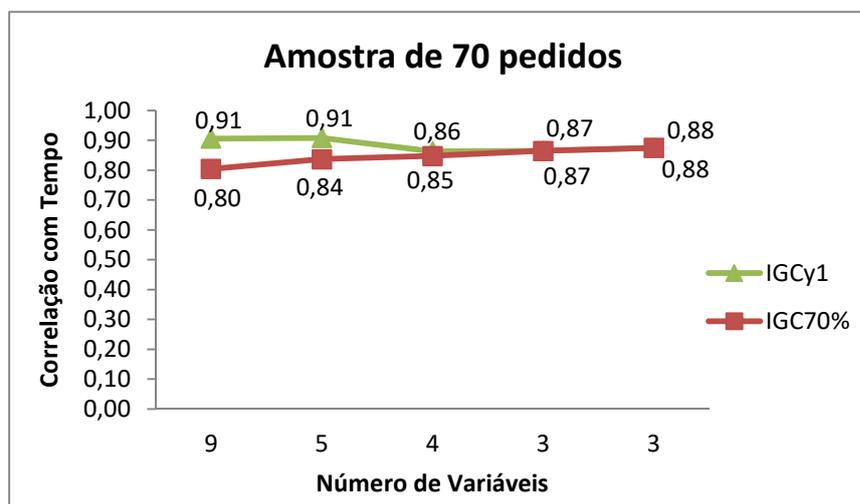


Figura 35 – Correlações do IGC e IGC_{Y1} com o tempo para amostras de 70 pedidos

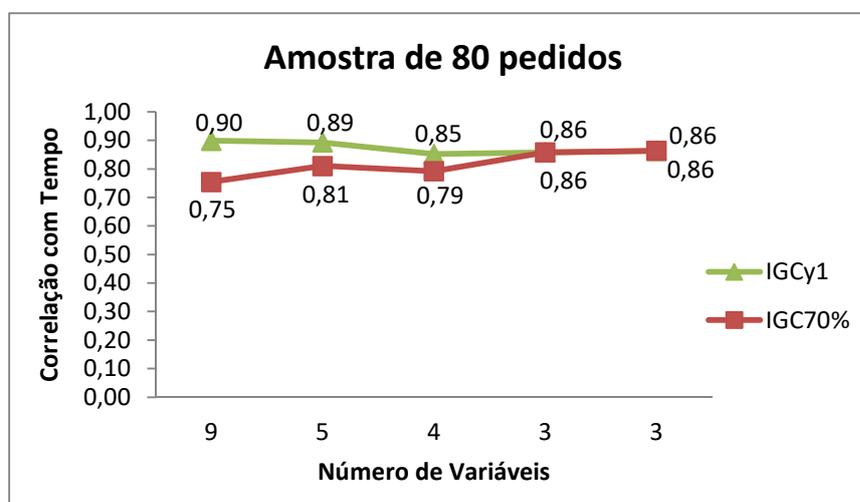


Figura 36 – Correlações do IGC e IGC_{Y1} com o tempo para amostras de 80 pedidos

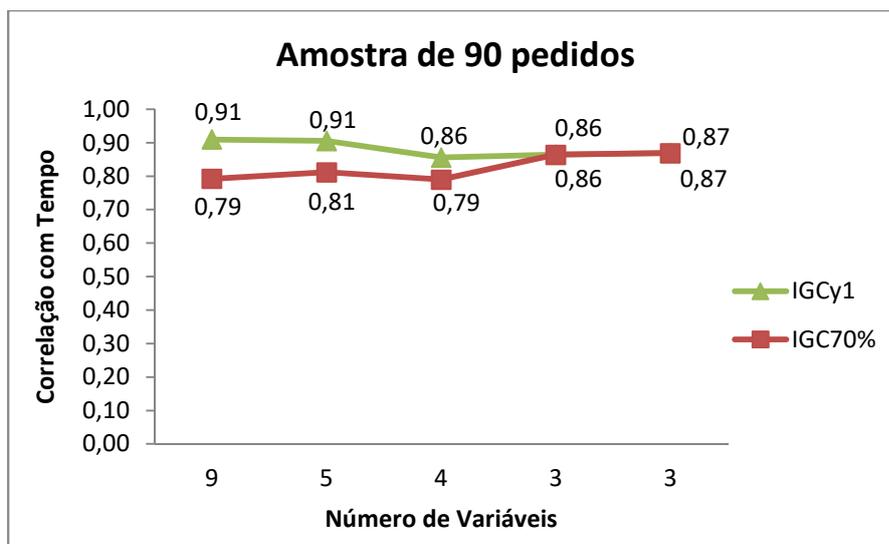


Figura 37 – Correlações do IGC e IGC_{Y1} com o tempo para amostras de 90 pedidos

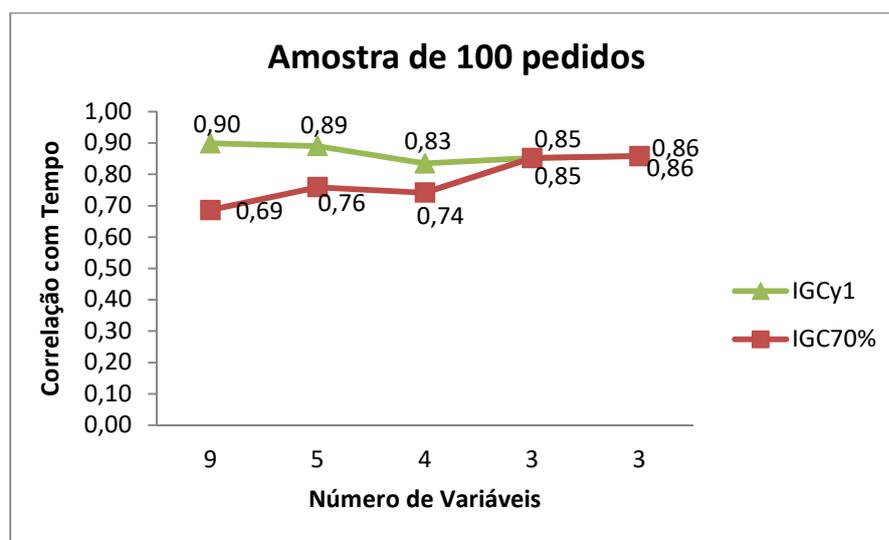


Figura 38 – Correlações do IGC e IGC_{Y1} com o tempo para amostras de 100 pedidos

A Figura 39 apresenta os gráficos das correlações do IGC_{Y1} com o tempo para todos os cinco casos e todas as amostras.

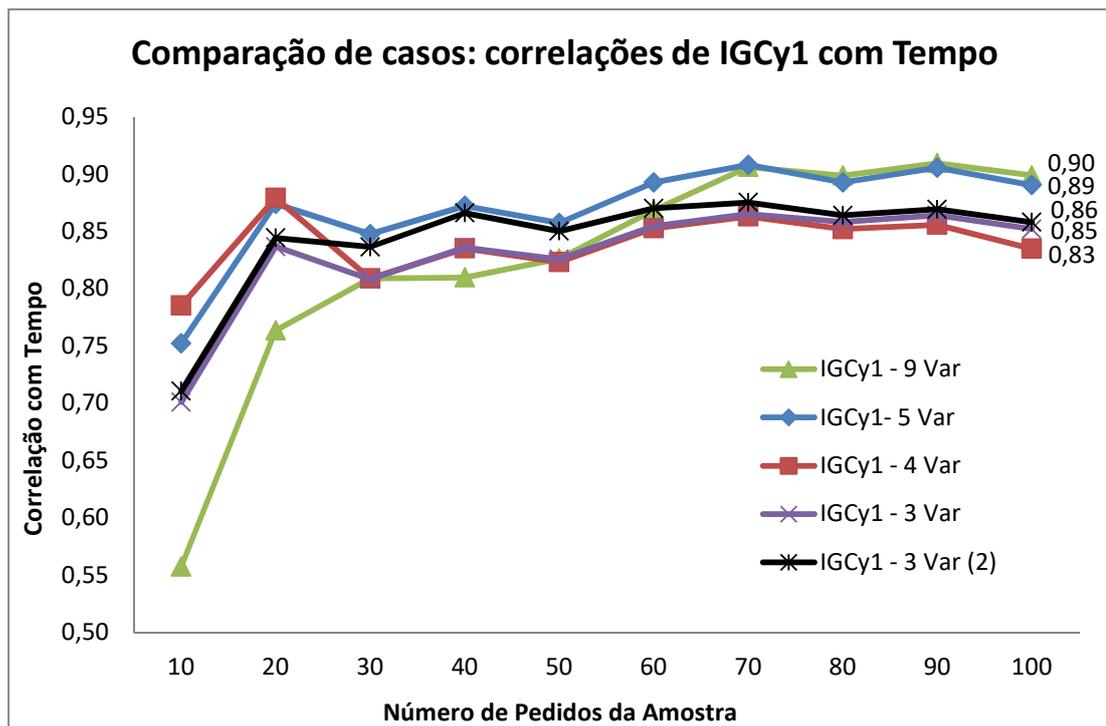


Figura 39 – Correlações do IG_{Cy1} com o tempo para todos os 5 casos e amostras de 10 a 100 pedidos

Da análise das Figuras 29 a 39, verifica-se que o perfil das curvas é bastante similar, com uma tendência de aumento de correlação do IGC com o tempo no início de todas elas, ou seja, quando o número de variáveis reduz de nove para cinco, sendo que a partir de cinco até três variáveis, as curvas se estabilizam mais rapidamente, atingem suas respectivas correlações “de estabilidade”. Estes resultados refletem as análises anteriores onde foi visto que quando se selecionam apenas as variáveis de volume direto, variando suas combinações, a tendência foi de obter correlações com tempo elevadas e estáveis com amostras a partir de somente vinte pedidos, ao passo que o caso com mais variáveis (nove), tende a estabilizar suas correlações a partir de uma amostra com cinquenta pedidos.

Destaque-se ainda que não houve diferenças tão significativas nas médias das correlações “de estabilidade” com tempo dos casos com nove até três variáveis e, sendo assim, embora todas as nove variáveis do estudo pareçam contribuir para a complexidade do pedido de patente, os resultados sugerem que, na prática, os casos com somente variáveis de volume direto já representam bem o tempo/esforço de exame.

Nota-se também que há um grande salto nas correlações com o tempo de todos os casos quando a amostra passa de dez para vinte pedidos de patente, algo que já era esperado,

pois os casos com somente dez pedidos não passaram em todos os testes e, conseqüentemente, tendem a apresentar menor eficiência na aplicação da ACP.

Por outro lado, é importante destacar que as correlações do $IGC_{70\%}$ com o tempo foram sempre menores ou iguais as do IGC_{Y1} , com uma tendência de igualdade quanto menor o número de variáveis selecionadas, sendo que para amostras de vinte ou mais pedidos e para quatro ou menos variáveis, foram sempre iguais. Destaque-se ainda que as correlações do IGC_{Y1} com o tempo permaneceram altas e quase constantes para quaisquer casos analisados com amostras a partir de vinte pedidos, mostrando um comportamento bastante estável independente do tamanho da amostra. Por conseguinte, para o problema em especial, os resultados convergem mostrando que o índice IGC_{Y1} é o mais adequado para a representação do tempo/esforço de exame.

Por fim, os resultados obtidos sugerem que o Caso 3 Var (2) é o de maior custo benefício para realização de testes práticos ainda mais específicos, seja por captar as influências das principais variáveis de volume direto de dados, seja pela maior simplicidade para obtenção e coleta de dados (pois não precisa separar reivindicações independentes e dependentes), ou mesmo pelas elevadas correlações de IGC e IGC_{Y1} com o tempo.

5.2.2.3 Enquadramento em Classes dos pedidos de patente da Amostra Padrão

Após verificar as correlações dos IGC dos casos selecionados com o tempo, foi realizado o enquadramento dos pedidos de patente nas cinco classes previstas: muito leve, leve, moderado, pesado e muito pesado.

5.2.2.3.1 Enquadramento dos pedidos de patente da Amostra Padrão por Tempo

Primeiramente os enquadramentos foram realizados com base na variável padronizada tempo de exame, sendo que todos os pedidos da amostra foram então enquadrados avaliando a média e os desvios padrão do tempo, de maneira similar às faixas já definidas na Tabela 2, ou seja:

- a) Se o tempo é maior que três ($\text{Tempo} > 3$, ou seja, é maior do que a média mais três desvios padrão) então o pedido é muito pesado;
- b) Se o tempo é maior do que um e menor ou igual a três ($1 < \text{Tempo} \leq 3$, ou seja, está no intervalo entre média mais um e média mais três desvios padrão) então o pedido é pesado;

c) Se o tempo está no intervalo entre menos um e um ($-1 \leq \text{Tempo} \leq 1$, ou seja, está no intervalo entre média mais ou menos um desvio padrão) então o pedido é moderado;

d) Se tempo está no intervalo entre menos três e menos um ($-3 \leq \text{IGC} < -1$, ou seja, está no intervalo entre média menos desvio padrão e média menos três desvios padrão) então o pedido é leve; e

e) Se o tempo é menor que menos três ($\text{Tempo} < -3$, ou seja, é menor do que a média menos três desvios padrão) então o pedido é muito leve.

Com isso foi obtida uma lista de enquadramentos de referência de todos os pedidos da amostra padrão com tempo. A Figura 40 mostra os tempos padronizados (eixo vertical) para todos os cem pedidos (eixo horizontal) da amostra padrão. As retas horizontais facilitam a verificação visual dos enquadramentos. Já a Figura 41 mostra a quantidade de enquadramentos por tempo nas classes.

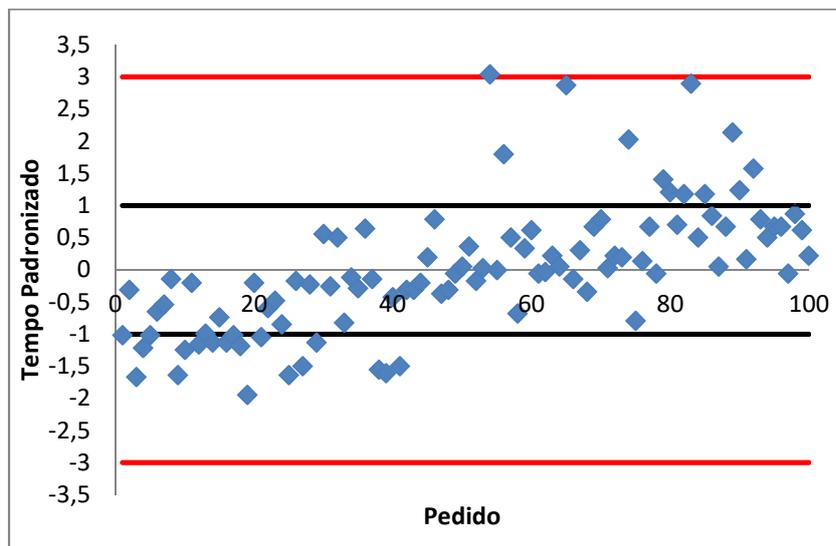


Figura 40 – Tempo padronizado dos pedidos

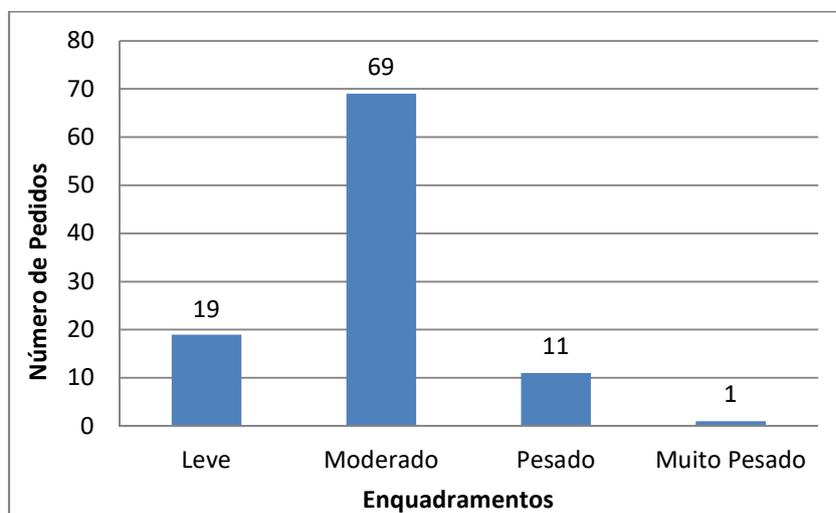


Figura 41 – Enquadramentos por tempo dos pedidos da amostra padrão

Da análise das Figuras 40 e 41, verifica-se que nenhum dos pedidos de patente da amostra foi enquadrado como muito leve e apenas um como muito pesado. A maioria dos pedidos foi enquadrada como moderado (69%), em segundo lugar como leve (19%), seguidos pelos pesados (11%). Este resultado se mostra coerente com os dados obtidos, já que a amostra padrão é relativamente homogênea (pedidos do mesmo tipo de exame - todos 6.21), sendo que a amostra da variável tempo apresentou distribuição estatística normal, além de coeficiente de variação moderado (21,61% - vide tabela 25).

As Tabelas 36 a 38 detalham os dados dos pedidos enquadrados como leves, pesados e muito pesado com relação ao tempo.

Tabela 36 – Pedidos de patente enquadrados como leves por tempo

Pedidos Leves no Enquadramento por Tempo											
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações	Total de Subclasses	Ano de Depósito	Inventores	Prioridades
1	6	4	3	13	2	13	15	4	11	1	1
3	25	2	7	34	1	4	5	1	10	2	0
4	11	4	2	17	1	6	7	3	10	1	1
5	19	2	11	32	1	2	3	1	11	2	1
9	6	2	0	8	1	6	7	1	12	1	1
10	9	2	2	13	1	2	3	1	12	1	1
12	6	3	6	15	1	9	10	1	12	1	1
14	22	2	6	30	1	5	6	2	12	3	1
16	12	2	9	23	1	6	7	1	12	5	0
17	10	3	2	15	2	10	12	1	12	2	0
18	11	2	5	18	1	8	9	1	12	3	1
19	10	1	1	12	1	0	1	2	12	2	1
21	8	3	6	17	4	3	7	1	12	5	1
25	4	1	2	7	1	4	5	1	12	2	1
27	16	1	5	22	1	3	4	1	12	1	1
29	16	2	8	26	2	4	6	3	12	1	1
38	12	1	4	17	1	2	3	1	12	3	1
39	9	1	3	13	1	2	3	1	12	1	1
41	4	1	1	6	2	5	7	1	12	1	1

Fonte: Elaboração Própria.

Tabela 37 – Pedidos de patente enquadrados como pesados por tempo

Pedidos Pesados no Enquadramento por Tempo											
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações	Total de Subclasses	Ano de Depósito	Inventores	Prioridades
56	40	4	19	63	2	11	13	1	12	1	1
65	88	8	14	110	2	12	14	1	12	4	4
74	47	4	6	57	2	15	17	4	12	1	3
79	32	8	6	46	3	20	23	1	12	1	1
80	33	5	4	42	3	15	18	1	10	1	0
82	23	6	4	33	4	17	21	4	9	1	1
83	80	6	18	104	2	13	15	1	10	2	0
85	40	5	7	52	4	8	12	2	12	1	1
89	67	7	17	91	3	8	11	1	12	6	1
90	16	8	4	28	3	21	24	1	12	2	1
92	19	6	4	29	5	28	33	2	12	2	2

Fonte: Elaboração Própria.

Tabela 38 – Pedido de patente enquadrado como muito pesado por tempo

Pedido Muito Pesado no Enquadramento por Tempo											
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações	Total de Subclasses	Ano de Depósito	Inventores	Prioridades
54	41	11	12	64	1	30	31	1	9	4	0

Fonte: Elaboração Própria.

Como pode ser observado das Tabelas 36 a 38, os enquadramentos com o tempo mostram um padrão específico quanto a três variáveis: número de páginas de reivindicações, número total de páginas e número total de reivindicações. O pedido 54 (vide tabela 38) apresentou valores bem altos nestas três variáveis, levando a um tempo de exame muito acima da média (22,67 horas) e, conseqüentemente, foi enquadrado como muito pesado. Pedidos com valores relativamente altos destas três variáveis, ou mesmo em pelo menos duas delas, levaram a tempos de exame razoavelmente acima da média e, conseqüentemente, foram enquadrados como pesados. Por outro lado, pedidos com valores baixos em duas ou três destas variáveis levaram tempos de exame razoavelmente abaixo da média sendo, conseqüentemente, enquadrados como leves. Estes resultados sugerem que estas três variáveis são as mais relevantes para uma avaliação do tempo/esforço de exame, de modo que reforçam que o Caso 3 Var (2) é o mais indicado para utilização prática e execução de futuros testes de redistribuição.

5.2.2.3.2 Enquadramento dos pedidos de patente da Amostra Padrão por IGC_{Y1}

De posse destes enquadramentos de referência por tempo obtidos no item anterior, processo análogo de enquadramento foi realizado considerando a média e desvios padrão do IGC_{Y1} ao invés do tempo. Para realização destes enquadramentos também será considerado o caso de melhor custo-benefício, ou seja, o Caso 3 Var (2) com a amostra padrão completa.

A Figura 42 mostra os IGC_{Y1} padronizados (eixo vertical) para todos os cem pedidos (eixo horizontal) da amostra padrão. As retas horizontais facilitam a verificação visual dos enquadramentos. Já a Figura 43 mostra a quantidade de enquadramentos por IGC_{Y1} nas classes.

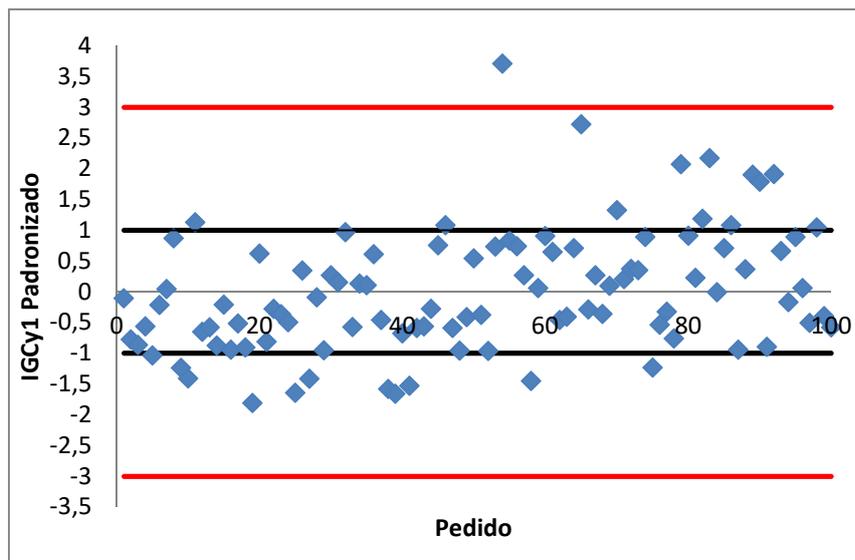


Figura 42 – IGC_{Y1} padronizado dos pedidos

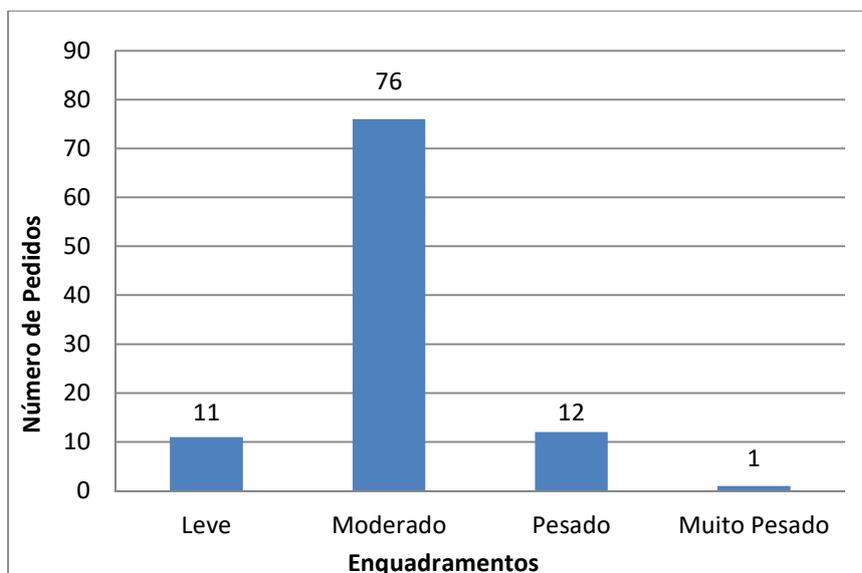


Figura 43 – Enquadramentos por IGC_{Y1} dos pedidos da amostra padrão

Da análise das Figuras 42 e 43, verifica-se que, de maneira similar aos enquadramentos por tempo, nenhum dos pedidos de patente da amostra foi considerado como muito leve e apenas um como muito pesado. A maioria dos pedidos foi enquadrada como moderado (76%), em segundo lugar como pesados (12%), seguidos pelos leves (11%).

Vale ressaltar novamente que este resultado se mostra coerente com os dados obtidos sendo que, assim como a amostra de tempos de exame, a amostra de dados de IGC_{Y1} também apresentou distribuição normal.

As Tabelas 39 e 41 detalham os dados dos pedidos enquadrados como leves, pesados e muito pesado com relação ao IGC_{Y1}. Já a Tabela 42 mostra os pedidos de patente em que houve diferença entre os enquadramentos por tempo e por IGC_{Y1}.

Tabela 39 – Pedidos de patente enquadrados como leves por IGC_{Y1}

Pedidos Leves no Enquadramento por IGC _{Y1}											
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações	Total de Subclasses	Ano de Depósito	Inventores	Prioridades
5	19	2	11	32	1	2	3	1	11	2	1
9	6	2	0	8	1	6	7	1	12	1	1
10	9	2	2	13	1	2	3	1	12	1	1
19	10	1	1	12	1	0	1	2	12	2	1
25	4	1	2	7	1	4	5	1	12	2	1
27	16	1	5	22	1	3	4	1	12	1	1
38	12	1	4	17	1	2	3	1	12	3	1
39	9	1	3	13	1	2	3	1	12	1	1
41	4	1	1	6	2	5	7	1	12	1	1
58	19	1	7	27	1	1	2	2	12	2	1
75	12	2	1	15	1	4	5	1	12	4	1

Fonte: Elaboração Própria.

Tabela 40 – Pedidos de patente enquadrados como pesados por IGC_{Y1}

Pedidos Pesados no Enquadramento por IGC _{Y1}											
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações	Total de Subclasses	Ano de Depósito	Inventores	Prioridades
11	24	6	7	37	2	17	19	3	12	2	1
46	38	6	11	55	2	11	13	1	12	4	1
65	88	8	14	110	2	12	14	1	12	4	4
70	27	8	7	42	2	11	13	1	12	7	1
79	32	8	6	46	3	20	23	1	12	1	1
82	23	6	4	33	4	17	21	4	9	1	1
83	80	6	18	104	2	13	15	1	10	2	0
86	24	6	1	31	2	18	20	1	12	2	1
89	67	7	17	91	3	8	11	1	12	6	1
90	16	8	4	28	3	21	24	1	12	2	1
92	19	6	4	29	5	28	33	2	12	2	2
98	26	6	4	36	2	16	18	1	12	1	1

Fonte: Elaboração Própria.

Tabela 41 – Pedido de patente enquadrado como muito pesado por IGC_{Y1}

Pedido Muito Pesado no Enquadramento por IGC _{Y1}											
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações	Total de Subclasses	Ano de Depósito	Inventores	Prioridades
54	41	11	12	64	1	30	31	1	9	4	0

Fonte: Elaboração Própria.

Tabela 42 – Pedidos de patente com diferença nos enquadramentos por tempo e IGC_{Y1}

Comparação de Enquadramentos por Tempo e IGC _{Y1}					
Pedido	Páginas de Reivindicações	Total de Páginas	Total de Reivindicações	Enquadramento por Tempo	Enquadramento por IGC _{Y1}
1	4	13	15	Leve	Moderado
3	2	34	5	Leve	Moderado
4	4	17	7	Leve	Moderado
11	6	37	19	Moderado	Pesado
12	3	15	10	Leve	Moderado
14	2	30	6	Leve	Moderado
16	2	23	7	Leve	Moderado
17	3	15	12	Leve	Moderado
18	2	18	9	Leve	Moderado
21	3	17	7	Leve	Moderado
29	2	26	6	Leve	Moderado
46	6	55	13	Moderado	Pesado
56	4	63	13	Pesado	Moderado
58	1	27	2	Moderado	Leve
70	8	42	13	Moderado	Pesado
74	4	57	17	Pesado	Moderado
75	2	15	5	Moderado	Leve
80	5	42	18	Pesado	Moderado
86	6	31	20	Moderado	Pesado
98	6	36	18	Moderado	Pesado

Fonte: Elaboração Própria.

Como pode ser observado das Tabelas 39 a 42, houve um total de vinte pedidos de patente com enquadramentos divergentes por tempo e IGC_{Y1}. Portanto, este resultado mostra que houve convergência no enquadramento de 80 dos 100 pedidos (80% do total), ou seja, uma grande similaridade.

Mais especificamente quanto às diferenças encontradas, verificam-se cinco enquadramentos adicionais pelo IGC_{Y1} na classe pesado (pedidos 11, 46, 70, 86 e 98). Como pode ser visto das Tabelas 40 e 42, tais pedidos apresentam número de páginas de reivindicações, número total de páginas e número total de reivindicações acima da média, apresentando perfil semelhante aos demais cinco pedidos pesados em que houve similaridade nos enquadramentos e, sendo assim, se justifica seu enquadramento pesado pelo IGC_{Y1}. Por outro lado, nota-se que os pedidos 46, 86 e 98 apresentaram valores de IGC_{Y1} próximos a um, ou seja, do limite de enquadramento entre as faixas moderado e pesado. Consequentemente, há dois fatores que possivelmente explicam este fenômeno: i) erros inerentes ao modelo matemático, os quais ainda que em pequena quantidade, tendem a ocorrer dependendo das variáveis, amostras e critérios adotados; ii) erros ou desvios de medição no tempo, o que poderia mover um enquadramento perto do limite das faixas.

Verificam-se também dez enquadramentos pelo IGC_{Y1} na classe moderado (pedidos 1, 3, 4, 12, 14, 16, 17, 18, 21 e 29) e com enquadramentos leves pelo tempo. Quanto aos pedidos 3, 14, 16, 18 e 29, verifica-se que apresentaram valores de IGC_{Y1} próximos a um, ou seja, do limite de enquadramento entre as faixas moderado e leve. Mais uma vez, tais divergências podem ser explicadas por dois fatores: i) erros inerentes ao modelo matemático, os quais ainda que em pequena quantidade, tendem a ocorrer dependendo das variáveis, amostras e critérios adotados; ii) erros ou desvios de medição no tempo, o que poderia mover um enquadramento perto do limite das faixas. Especificamente com relação ao pedido 1, nota-se que ele apresentou tanto o número de páginas de reivindicações quanto número total de reivindicações acima da média, o que reforça seu enquadramento como moderado pelo modelo. Além disso, os pedidos 1, 17 e 21 apresentaram tempos padronizados bem próximos de menos um (respectivamente -1,02, -1,02 e -1,04), ou seja, bem perto do limite de enquadramento entre as faixas leve e moderado. Nestes casos a tendência é de que desvios na medição do tempo possam ter gerado a divergência. Já os pedidos 4 e 12 possuem poucas páginas e número de páginas de reivindicações perto da média. Em princípio, há uma tendência a serem considerados leves, a depender do peso que se dê especificamente ao quadro reivindicatório ou ao total de páginas. Embora não haja certeza, a tendência é que nesses casos erros inerentes ao modelo matemático tenham gerado as divergências.

Adicionalmente, verificam-se três enquadramentos pelo IGC_{Y1} na classe moderado (pedidos 56, 74 e 80) e com enquadramentos pesados pelo tempo. Os pedidos 74 e 80 apresentaram valores de IGC_{Y1} próximos a um, ou seja, do limite de enquadramento entre as faixas moderado e leve. Mais uma vez, tais divergências podem ser explicadas por dois fatores: i) erros inerentes ao modelo matemático, os quais ainda que em pequena quantidade, tendem a ocorrer dependendo das variáveis, amostras e critérios adotados; ii) erros ou desvios de medição no tempo, o que poderia mover um enquadramento perto do limite das faixas. Especificamente com relação ao pedido 56, verifica-se que é um pedido com muitas páginas, todavia, com número de páginas de reivindicações e número total de reivindicações pouco acima da média. A depender do procedimento específico de exame realizado, necessidade ou não de um maior entendimento do relatório e das figuras, o tempo pode levar a um enquadramento moderado ou pesado. Sendo assim, é um tipo de pedido difícil de ser enquadrado *a priori* e, neste caso, o modelo acabou atuando de maneira mais conservadora, enquadrando como moderado.

Por fim, verificam-se dois enquadramentos pelo IGC_{Y1} na classe leve (pedidos 58 e 75) e com enquadramentos moderados pelo tempo. Tais pedidos apresentam número de páginas de reivindicações (respectivamente uma e duas) e número total de reivindicações (respectivamente duas e cinco) bem abaixo da média. O pedido 58 possui ainda um número total de páginas abaixo (mas próximo) da média e o pedido 75 bem abaixo da média. Em uma avaliação *a priori*, o enquadramento leve do modelo faz todo o sentido nestes casos. Sendo assim, embora não haja certeza, a tendência é de que desvios na medição do tempo possam ter gerado tais divergências.

Além das análises detalhadas anteriores, foram ainda avaliados os acertos de enquadramentos para todos os cinco casos de estudo e para todas as amostras de dez a cem pedidos. A Figura 44 a seguir apresenta estes resultados.

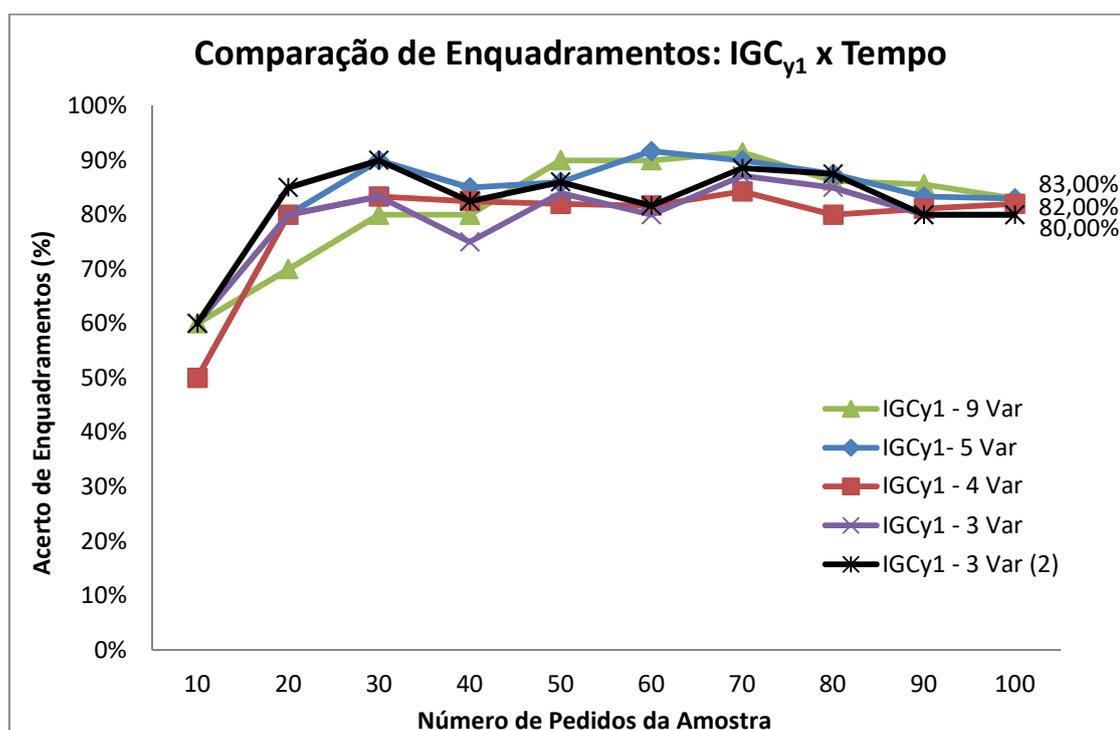


Figura 44 – Correlações do IGC_{Y1} com o tempo para todos os 5 casos e amostras de 10 a 100 pedidos

Da análise da Figura 44 verifica-se que os quatro casos considerando apenas as variáveis de volume direto (casos 5 Var, 4 Var, 3 Var e 3 Var (2)) apresentam uma boa estabilidade dessa quantidade de acertos de enquadramento já a partir de amostras com 20 pedidos de patente. Por outro lado, o caso completo (caso 9 Var - com todas as nove variáveis) também atinge a estabilidade, porém, somente a partir de amostras com 40 a 50

pedidos de patente. Tais resultados estão dentro do esperado dado que, em termos práticos, o ideal são amostras com total de pedidos de pelo menos cinco vezes o número de variáveis do caso. Além disso, todos os casos apresentam boa similaridade entre os enquadramentos por IGC_{Y1} e por tempo de exame, oscilando na faixa de 80 a 90%, com média geral de 81%.

Em suma, verifica-se que o modelo consegue representar de maneira bastante satisfatória o esforço/tempo de exame, sendo que para casos no limiar dos critérios adotados ocorrem poucas divergências e, nestes casos, as divergências ocorrem apenas em classes de enquadramento adjacentes. Ou seja, ocorrem divergências pontuais, não grosseiras e justificáveis dadas as limitações inerentes a um modelo e pesquisa deste tipo.

De posse do modelo matemático e de critérios de enquadramento testados com as correlações com o tempo, o próximo passo será apresentar a seguir um teste de redistribuição completo.

5.3 TESTE DE REDISTRIBUIÇÃO – ÁREA DE ELETRICIDADE

5.3.1 Amostra Final de Redistribuição

Para a composição da amostra final de redistribuição, foram selecionados dez pedidos de patente de cada um dos dez examinadores em análise (um examinador a menos do que na amostra de teste inicial, pois um entrou de licença), perfazendo um total de cem (100) pedidos de patente a serem avaliados. Foram seguidos estritamente os passos da metodologia proposta no que concerne à seleção do marco temporal e demais premissas já definidas quanto ao arcabouço documental a ser analisado. Todavia, como neste caso o objetivo era obter uma amostra para aplicação do modelo validado com a amostra padrão com tempo (nossa referência), foram obtidas todas as variáveis de interesse de primeiros exames já publicados, coletados entre maio e julho de 2020, sendo que todos os pedidos da amostra também são do tipo 6.21, ou seja, com aproveitamento de buscas de escritórios internacionais.

A Tabela 43 mostra, para o conjunto total de 100 pedidos, um quadro geral das principais variáveis obtidas e os resultados dos cálculos de estatísticas básicas tais como: média, desvio padrão, mediana, moda, mínimo, máximo e coeficiente de variação.

Tabela 43 – Estatísticas Básicas da Amostra Final de Redistribuição

	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Reivindicações Independentes	Reivindicações Dependentes	Total de Subclasses	Ano de Depósito	Inventores	Prioridades
Média	21,15	3,91	6,85	1,90	9,28	1,52	11,35	2,55	1,06
Desvio Padrão	14,41	1,83	6,03	1,13	5,00	0,78	0,99	1,66	0,80
Mediana	18,00	4,00	5,00	2,00	8,50	1,00	11,00	2,00	1,00
Moda	16	4	3	1	7	1	12	1	1
Mínimo	4	1	0	1	0	1	8	1	0
Máximo	102	12	36	9	23	5	13	10	8
Coefficiente de Variação	68,12%	46,85%	87,97%	59,61%	53,87%	51,63%	8,71%	65,09%	75,64%

Fonte: Elaboração Própria.

Da análise da Tabela 43, verifica-se que, para quase todas as variáveis do pedido a serem consideradas, os desvios padrão (e variância, pois é o quadrado do desvio) e coeficientes de variação (relação entre o desvio padrão e a média) ainda são altos, embora menores do que os da amostra de teste inicial. A exceção é o ano de depósito que apresenta coeficiente de variação bem baixo (8,71%) e moda igual a 12, evidenciando que a distribuição original para os examinadores ficou concentrada em poucos anos de depósito, principalmente no ano de 2012. Adicionalmente, modas, médias e medianas das variáveis são distintas entre si, sendo que os valores mínimos e máximos são bastante diferentes. Tais fatos mais uma vez sugerem que alta amplitude e boa dispersão são características intrínsecas dos pedidos analisados, ou seja, há uma variabilidade razoável entre eles mesmo considerando uma amostra menor e com pedidos de mesmo tipo (6.21 com aproveitamento de buscas), sendo que, conforme desejado, o perfil dos dados é parecido com o da amostra padrão com tempo.

5.3.2 Execução da Análise de Componentes Principais: Caso 3 Var (2)

O caso escolhido para simulação da redistribuição foi o Caso 3 Var (2), dado que este caso foi o identificado como de melhor custo-benefício nos testes de validação com a amostra padrão com tempo. Todavia, agora os testes foram realizados com os dados da amostra final de redistribuição completa (incluindo os cem pedidos de patente coletados).

De maneira análoga a todos os casos anteriores, foram obtidos os autovalores (λ_1, λ_2 e λ_3), a matriz V de autovetores associados (escores das componentes), as componentes principais e suas respectivas contribuições de variâncias, além da contribuição acumulada das componentes (% de variância acumulada). A Tabela 44 mostra os principais resultados.

Tabela 44 – Autovalores e Contribuições de Variância das Componentes Principais

Componente Principal	Autovalores	Variância	Variância Acumulada
Y1	1,82	61%	61%
Y2	0,94	31%	92%
Y3	0,24	8%	100%

Fonte: Elaboração Própria.

Da análise da Tabela 44, verifica-se que no caso de utilização do critério considerando autovalores maiores do que um, ou mesmo o critério de variância acumulada (acima de 60%), seleccionaríamos somente a primeira componente para a determinação do IGC, ou seja, $IGC_{\lambda>1} = IGC_{Y1}$.

A Tabela 45 mostra os escores das duas primeiras componentes principais para cada uma das variáveis consideradas do caso 3 var (2) com 100 pedidos.

Tabela 45 – Escores das Componentes Principais de cada variável

Variáveis	Coeficiente de Ponderação	
	Y1	Y2
Total de Páginas	0,4492	-0,8879
Total de Reivindicações	0,5649	0,4576
Páginas de Reivindicações	0,6922	0,0469

Fonte: Elaboração Própria.

Na Tabela 45 foram hachurados na cor cinza os valores considerados significativos dos fatores para cada uma das variáveis (próximos ou acima de 0,4). Em uma primeira análise, nota-se que a componente Y1 é bastante destacada das demais, sendo responsável por explicar 61% de toda a variância. Neste caso, ambas as componentes Y1 e Y2 estão associadas às principais variáveis diretamente relacionadas ao volume de dados (total de páginas, páginas de reivindicações, além de número de reivindicações). Este resultado é bem semelhante aos demais casos 3 var (2) já executados.

5.3.3 Enquadramento dos Pedidos de Patente da Amostra Final de Redistribuição por IGC_{Y1}

A Figura 45 mostra os IGC_{Y1} padronizados (eixo vertical) para todos os cem pedidos (eixo horizontal) da amostra. As retas horizontais facilitam a verificação visual dos enquadramentos. Já a Figura 46 mostra a quantidade de enquadramentos por IGC_{Y1} nas classes.

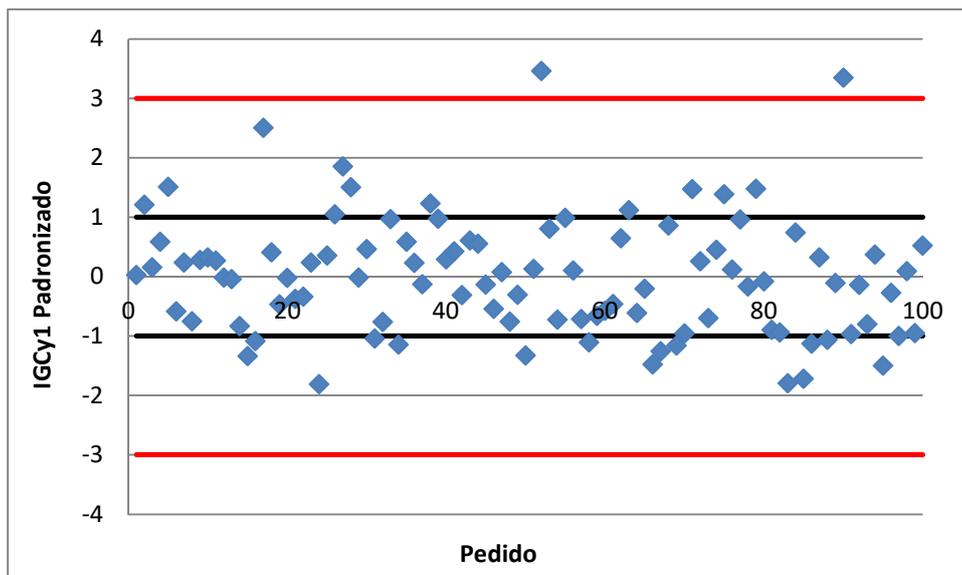


Figura 45 – IGC_{Y1} padronizado dos pedidos

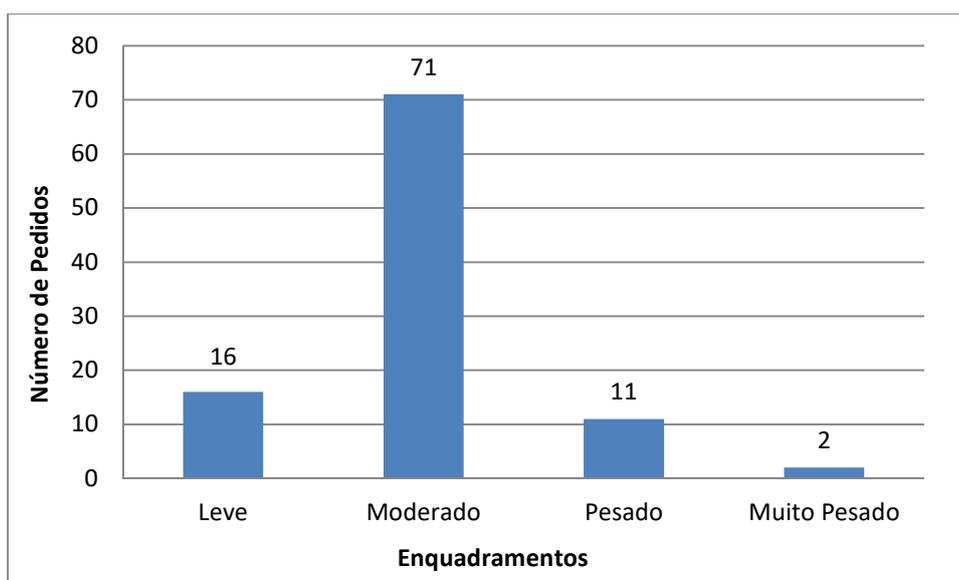


Figura 46 – Enquadramentos por IGC_{Y1} dos pedidos da amostra padrão

Da análise das Figuras 45 e 46, verifica-se que nenhum dos pedidos de patente da amostra foi enquadrado como muito leve. A maioria dos pedidos foi enquadrada como moderado (71%), em segundo lugar como leve (16%), seguidos pelos pesados (11%) e, por fim, os muito pesados, que foram apenas dois (2%). Vale ressaltar novamente que este resultado se mostra coerente com o dos testes anteriores e, assim como no caso da amostra padrão com tempo, neste caso a amostra de dados de IGC_{Y1} também apresentou distribuição normal.

As Tabelas 46 a 48 detalham os dados dos pedidos enquadrados como muito pesados, pesados e leves com relação ao IGC_{Y1}.

Tabela 46 – Pedidos de patente enquadrados como muito pesados por IGC_{Y1}

Pedidos Muito Pesados no Enquadramento por IGC _{Y1}							
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
52	69	11	31	111	9	4	13
90	24	12	6	42	4	19	23

Fonte: Elaboração Própria.

Tabela 47 – Pedidos de patente enquadrados como pesados por IGC_{Y1}

Pedidos Pesados no Enquadramento por IGC _{Y1}							
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
2	14	6	4	24	1	20	21
5	25	7	16	48	3	13	16
17	42	8	4	54	3	21	24
26	22	6	0	28	3	15	18
27	26	7	8	41	3	19	22
28	19	5	7	31	4	23	27
38	32	5	5	42	3	18	21
63	38	6	11	55	2	11	13
71	39	7	9	55	2	12	14
75	22	6	2	30	2	20	22
79	102	4	36	142	2	4	6

Fonte: Elaboração Própria.

Tabela 48 – Pedidos de patente enquadrados como leves por IGC_{Y1}

Pedidos Leves no Enquadramento por IGC _{Y1}							
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
15	9	2	15	26	1	1	2
16	10	2	6	18	1	6	7
24	7	1	2	10	1	2	3
31	15	2	13	30	1	4	5
34	6	2	2	10	1	7	8
50	12	2	8	22	1	2	3
58	12	2	3	17	1	6	7
66	9	2	2	13	1	2	3
67	6	2	0	8	1	6	7
69	19	2	11	32	1	2	3
83	7	1	3	11	1	2	3
85	16	1	3	20	1	1	2
86	7	2	2	11	2	6	8
88	19	2	3	24	1	5	6
95	16	2	3	21	1	0	1
97	10	2	2	14	1	8	9

Fonte: Elaboração Própria.

Como pode ser observado das Tabelas 46 a 48, conforme esperado os enquadramentos pelo IGC_{Y1} mostram um padrão específico quanto às três variáveis do caso simulado: número de páginas de reivindicações, número total de páginas e número total de reivindicações.

De uma análise específica da Tabela 46, verifica-se que os pedidos 52 e 90 apresentaram uma quantidade de páginas de reivindicações (11 e 12) muito acima da média (3,8), além de um total de reivindicações (13 e 23) também acima da média (11). No caso do número total de páginas, o pedido 52 também apresentou valor (111) muito acima da média (32,4) e o pedido 90 um pouco acima (42) da média. Conseqüentemente, ambos foram enquadrados como muito pesados.

Considerações análogas às anteriores podem ser feitas para os pedidos enquadrados como pesados, conforme ilustra a Tabela 47. Dentre os pedidos pesados, cabe um destaque para o pedido 79, que apresentou o maior número total de páginas (142). Todavia, ele apresenta quatro (4) páginas de quadro reivindicatório, valor quase que idêntico à média geral (3,8), e ainda apresenta somente seis (6) reivindicações, valor abaixo da média (10). Verifica-se que, apesar do pedido apresentar características de muito esforço de leitura (etapa 1 do exame), a análise das reivindicações/patenteabilidade (etapa 3 de exame) estaria em nível considerado moderado. Conseqüentemente, o modelo sugeriu um enquadramento como pesado, ou seja, um meio termo.

Por fim, da análise da Tabela 48 também pode-se observar um padrão para as três variáveis avaliadas. Todos os pedidos apresentaram uma ou duas páginas de reivindicações, valor abaixo da média. O mesmo pode ser dito quanto ao número total de reivindicações e número total de páginas que, embora alguns apresentem valores próximos da média, todos foram abaixo da média. Todavia, os valores não foram tão abaixo da média a ponto de gerar algum enquadramento muito leve. Especificamente com relação aos pedidos 24 e 83, verifica-se que eles são bem pequenos, com apenas uma página de reivindicação, total de somente dez e onze páginas, e somente três reivindicações, sendo uma delas independente. Nestes casos, pela lógica os pedidos poderiam ser enquadrados como muito leves ao invés de leves, evidenciando que, para situações em que as três variáveis apresentem coeficiente de variação muito alto, o modelo pode apresentar limitações para enquadrar pedidos muito leves. Em situações de dúvidas para enquadramentos muito leves, é sugerida uma análise complementar do percentil 5 da amostra, ou seja, verificar os pedidos que, para as três variáveis, apresentem valores abaixo dos 5% menores. Os percentis 5 para as variáveis número de páginas de reivindicações, número total de páginas e número total de reivindicações foram,

respectivamente, iguais a 2, 11 e 3. Verifica-se que ambos os pedidos estão no limite do percentil 5 para duas das três variáveis, sendo que as páginas do quadro reivindicatório estão abaixo. Conseqüentemente, embora não seja indiscutível, tais pedidos poderiam ser enquadrados como muito leves e a análise de percentil 5 pode auxiliar na decisão. Sendo assim, é importante haver possibilidade da realização de um ajuste fino pelo gestor responsável ou, em caso de utilização do modelo para distribuição automática, haver possibilidade de interação entre o examinador e o gestor visando a possíveis alterações consensuais.

5.3.4 Análise das Subclasses da CIP da Amostra Final de Redistribuição

5.3.4.1 ZAE Original dos Examinadores da Amostra Final de Redistribuição

Na amostra de final de redistribuição foi encontrado um número total de 32 subclasses principais ao longo de todos os 100 pedidos de patentes analisados, um número que mais uma vez mostra uma boa dispersão tecnológica da área de eletricidade.

Conforme já verificado quando da análise da amostra de teste inicial, as ZAE de referência dos examinadores estão definidas conforme a tabela 15. A Tabela 49 a seguir mostra a configuração das zonas de atuação de cada um dos 10 examinadores por subclasses da CIP para a distribuição original da amostra de redistribuição. Destaca-se que a área cinza corresponde às subclasses das Zonas de Atuação Específicas (ZAE) dos examinadores.

Tabela 49 – Zonas de Atuação dos Examinadores por Subclasses da CIP

ZAE original da Amostra Final de Redistribuição																			
1		2		3		4		5		6		7		8		9		10	
Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%
H01H	60,00%	H01R	50,00%	H02P	20,00%	H02K	40,00%	G01R	10,00%	H01R	20,00%	B60T	20,00%	G07F	10,00%	H02G	20,00%	B60L	30,00%
H02G	20,00%	G01R	30,00%	G01R	10,00%	H01F	20,00%	F02D	60,00%	F16H	20,00%	H01H	30,00%	G03G	30,00%	H02B	10,00%	H03K	30,00%
H01T	10,00%	H01H	10,00%	H05B	10,00%	H02P	20,00%	B60R	30,00%	F02D	10,00%	H01F	10,00%	G05B	20,00%	B60T	30,00%	G05B	10,00%
H01C	10,00%	B60L	10,00%	H02J	30,00%	H02M	20,00%			G05B	10,00%	H02M	10,00%	B61L	20,00%	F16H	10,00%	H03F	10,00%
				H02G	20,00%					B23K	10,00%	H02H	10,00%	B62D	10,00%	H01H	10,00%	F02N	10,00%
				H01G	10,00%					F02M	10,00%	H02G	10,00%	A61M	10,00%			B60T	10,00%
				H01F	10,00%					H02N	10,00%	B61L	10,00%						
										H01F	10,00%								

Fonte: Elaboração Própria

Da análise da Tabela 49, verifica-se que as configurações de distribuição nas ZAE de cada um dos examinadores são bem variáveis. Enquanto os examinadores 6 e 10 não apresentam nenhum dos seus dez pedidos distribuídos dentro das ZAE de referência, os examinadores 5 e 8 apresentam um, o examinador 7 apresenta dois, o examinador 9 apresenta três, os examinadores 3 e 4 apresentam quatro, o examinador 2 apresenta cinco e o

examinador 1 apresenta oito. Tal fato mostra que, mesmo com uma tentativa de distribuição de carga atual utilizando como base a especialidade/áreas de conhecimento dos examinadores, caso a amostra não contenha pedidos suficientes ou mesmo essa distribuição não esteja sistematizada/automatizada de forma a auxiliar os gestores, podem ocorrer distorções indesejáveis.

5.3.4.2 Configuração Original dos Pedidos

A Tabela 50 mostra as medianas das variáveis dos pedidos com os dados separados por examinador. Nesta tabela foram hachurados: na cor cinza claro as variáveis dos pedidos de examinadores com mediana abaixo da mediana geral dos dados; e na cor cinza escuro as variáveis dos pedidos dos examinadores com mediana acima da mediana geral dos dados.

Tabela 50 – Medianas da Amostra de Final de Redistribuição separadas por examinador

Medianas							
Examinador	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
1	14,5	4,5	5	24,5	2	12	14
2	16	3	6,5	25,5	1	10	11,5
3	19,5	4,5	5	29,5	2	11	13
4	16	4	5	26,5	2	12,5	14,5
5	16,5	3,5	5,5	27	2	6,5	8,5
6	17	3,5	6,5	28	2	6,5	10,5
7	19	3	3,5	29	1	7	8,5
8	19	4	3,5	29,5	2	11,5	13,5
9	20	2	5	28	1	6	7,5
10	17,5	3	4	25	1	5	7,5
Divisão	18,0	4,0	5,0	27,5	2,0	8,5	10,0

Fonte: Elaboração Própria.

Da análise da Tabela 50, verifica-se que praticamente todos os examinadores apresentam, para as três variáveis em análise, medianas tanto abaixo quanto acima da mediana geral dos dados da divisão. Nota-se, ainda, que os examinadores 9 e 10 apresentam medianas do número de páginas de reivindicações e total de reivindicações razoavelmente abaixo das respectivas medianas gerais da divisão, sugerindo que podem estar com carga um pouco leve em relação aos demais. Já o examinador 3, apresenta medianas das três variáveis um pouco acima da mediana geral da divisão, sugerindo que possa estar levemente sobrecarregado em relação aos demais. Em suma, *a priori* o balanceamento parece estar bom, com pequenos desvios a tentar corrigir ou minimizar com uma nova redistribuição.

5.3.4.3 Redistribuição dos Pedidos

5.3.4.3.1 Nova Configuração das ZAE dos Examinadores

A Tabela 51 mostra a nova configuração das ZAE dos examinadores após ser realizada uma distribuição com a nova lógica desenvolvida. Cabe destacar que, na nova distribuição, o número de pedidos distribuídos para cada examinador permaneceu o mesmo que na distribuição original, ou seja, cada um com um total de 10 pedidos.

Tabela 51 – ZAE dos examinadores após a nova distribuição

Novas ZAE da Amostra Final de Redistribuição - IGC ₁₁																				
1		2		3		4		5		6		7		8		9		10		
Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	
H01H	50,00%	H01H	30,00%	H02P	20,00%	H02K	30,00%	H02M	30,00%	H02J	30,00%	B60T	30,00%	G07F	10,00%	H01R	50,00%	G01R	30,00%	
H02G	20,00%	H01R	20,00%	G01R	10,00%	B60L	30,00%	G01R	10,00%	H02P	20,00%	F02D	20,00%	B60T	20,00%	H02G	40,00%	H02B	10,00%	
F16H	20,00%	B60R	10,00%	H05B	10,00%	G03G	10,00%	G05B	10,00%	H02H	20,00%	B51L	20,00%	F02D	20,00%	B60L	10,00%	H01C	20,00%	
F02M	10,00%	H02N	10,00%	H05K	10,00%	F02D	10,00%	F02D	10,00%	H01H	10,00%	H01F	10,00%	H01F	20,00%			F02D	10,00%	
		H01G	10,00%	H02K	10,00%	F16H	10,00%	B60T	10,00%	G05B	10,00%	B60R	10,00%	H01T	10,00%			B60R	10,00%	
		F16H	10,00%	H01F	10,00%	A61M	10,00%	H01F	10,00%	H05K	10,00%	G03G	10,00%	H01H	10,00%			G03G	10,00%	
		G05B	10,00%	F02N	10,00%			B61L	10,00%					G05B	10,00%					
				F16H	10,00%			H03F	10,00%											
				B62D	10,00%														H05K	10,00%

Fonte: Elaboração Própria

A Tabela 52 mostra uma comparação das porcentagens de pedidos dentro das ZAE em relação ao total de pedidos distribuídos para cada examinador nas distribuições nova e original.

Tabela 52 – Porcentagem de pedidos dentro das ZAE antes e após a nova distribuição

Examinador	1	2	3	4	5	6	7	8	9	10
Nova ZAE	70,00%	50,00%	40,00%	60,00%	40,00%	50,00%	30,00%	10,00%	90,00%	40,00%
ZAE Original	80,00%	50,00%	40,00%	40,00%	10,00%	0,00%	20,00%	10,00%	30,00%	0,00%

Fonte: Elaboração Própria

Da análise das Tabelas 51 e 52, verifica-se que houve um nítido maior equilíbrio na concentração de exame dentro das ZAE de cada um dos examinadores. Ressalte-se que no caso de seis dos dez examinadores houve um aumento do número de pedidos distribuídos a eles e pertencentes às suas próprias ZAE, com destaque para os examinadores 5, 9 e 10, com aumentos bem expressivos. Já os examinadores 2, 3 e 8 permaneceram com as mesmas concentrações da distribuição original, e somente o examinador 1 reduziu um pedido, passando de 80% para 70%. Destaque-se que somente o examinador 8 se manteve com uma pequena concentração (10%) de exame dentro da sua ZAE, o que talvez possa ser explicado pelo fato de este ser um examinador mais versátil da divisão, que não tem uma ZAE enxuta e

bem definida. Além disso, vale destacar que, se na distribuição original tínhamos dois examinadores sem nenhum pedido distribuído em suas ZAE, com a nova distribuição todos os examinadores tiveram pelo menos um dos dez pedidos distribuídos em suas respectivas ZAE. Sendo assim, os resultados evidenciam que esta nova configuração contribui para um maior equilíbrio da distribuição, para que os examinadores trabalhem dentro das suas áreas de formação e conhecimento específicas e, conseqüentemente, para a sua eficiência e motivação.

5.3.4.3.2 Nova Configuração dos Pedidos Redistribuídos

A Tabela 53 mostra as medianas das variáveis dos pedidos redistribuídos com os dados separados por examinador. Nesta tabela foram hachurados: na cor cinza claro as variáveis dos pedidos de examinadores com mediana abaixo da mediana geral dos dados; e na cor cinza escuro as variáveis dos pedidos dos examinadores com mediana acima da mediana geral dos dados.

Tabela 53 – Medianas da Amostra de Final de Redistribuição separadas por examinador

Medianas							
Examinador	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
1	17,5	4	4	26,5	2	10	11,5
2	19,5	4	7	31	2	6,5	9
3	15,5	4	3	23	2	9	10,5
4	15	3,5	4	21,5	2	7	9
5	22,5	5	7	31	2	9	11
6	19	4	5	29,5	2	7	10
7	20,5	4	4	31	1	8	9
8	17	3	6	27	1,5	8,5	10
9	16,5	3,5	6,5	25	1	10,5	12
10	16,5	4	5,5	25,5	1	9	10,5
Geral	18	4	5	27,5	2	8,5	10

Fonte: Elaboração Própria.

Da análise da Tabela 53, verifica-se que ainda há algumas variáveis de pedidos dos examinadores com mediana tanto abaixo quanto acima da mediana geral dos dados da divisão. Todavia, nota-se que agora estes desvios foram reduzidos, principalmente com relação a variável número de páginas de reivindicações. O único examinador que apresentou medianas das três variáveis abaixo da mediana geral foi o de número 4, mas ainda assim com valores próximos aos da mediana geral. Por outro lado, o examinador 5 foi o único que apresentou medianas das três variáveis acima da mediana geral, sendo duas delas próximas da mediana geral, sugerindo que possa estar levemente sobrecarregado em relação aos demais. Sendo assim, *a priori* não foi identificado um examinador com tendências claras de carga leve

ou sobrecarga com relação aos demais. Estes resultados mostram que, além de melhorar as suas ZAE, a nova redistribuição teve um efeito positivo no ajuste fino das medianas das amostras dos examinadores, algo que será quantificado por meio do cálculo do IBD e avaliado por testes de hipótese.

5.3.4.3.3 Índice de Balanceamento da Distribuição (IBD) antes e após a Redistribuição

Para complementar o ciclo da metodologia e realizar uma última medida de comparação quantitativa entre as distribuições, foram calculados os índices IBD da distribuição original da amostra e da sua redistribuição. Para realização dos cálculos foram utilizadas as medianas constantes da tabela 50 (distribuição original) e da tabela 53 (redistribuição).

Por meio da equação (16), foram então obtidos IBD igual a 0,83 para o caso original e IBD igual a 0,90 para a redistribuição, ou seja, houve um aumento de 8,43% do IBD com nova distribuição, demonstrando que as medianas dos pedidos dos examinadores após a redistribuição se aproximam mais da mediana geral da divisão. Tal fato corrobora que com a nova distribuição temos uma tendência de maior equilíbrio quanto ao volume de dados e tempo/esforço dos pedidos distribuídos aos examinadores.

5.3.4.3.4 Testes de Hipótese

Para uma última análise complementar dos resultados da nova distribuição, foram executados testes de hipótese de Mann-Whitney de modo a comparar as distribuições estatísticas e medianas das amostras de pedidos distribuídos a cada examinador com a distribuição estatística e mediana da amostra geral da divisão (Gibbons, 1971; Noether, 1976; Viali, 1976). Os testes foram aplicados para as amostras das três variáveis de interesse (número de páginas de reivindicações, número total de reivindicações e número total de páginas) dos 10 examinadores. Vale destacar que para a realização dos testes foi considerado um nível de significância de 0,01 e, sendo assim, o teste em questão tem como resultado positivo, ou seja, as medianas das amostras podem ser consideradas aproximadamente iguais, quando o p-valor encontrado for maior do que 0,01. Destaque-se ainda que, quanto mais perto de um for o p-valor, teoricamente melhor.

A Tabela 54 mostra os resultados de p-valor dos testes de Mann-Whitney realizados para cada uma das três variáveis e cada um dos dez examinadores. Para obtenção dos

resultados foi utilizada calculadora gratuita disponível na internet (<https://www.socscistatistics.com/tests/mannwhitney/default2.aspx>).

Tabela 54 – Resultados de p-valor dos Testes de Mann-Whitney

Examinador	p-valor dos testes de Mann-Whitney		
	Páginas de Reivindicações	Total de Reivindicações	Total de Páginas
1	0,749	0,704	0,529
2	0,952	0,787	0,624
3	0,772	0,596	0,219
4	0,242	0,390	0,197
5	0,276	0,904	0,204
6	0,582	0,865	0,424
7	0,726	0,749	0,631
8	0,447	0,984	0,841
9	0,857	0,490	0,728
10	0,772	0,905	0,803

Fonte: Elaboração Própria.

Da análise da Tabela 54, verifica-se que as amostras de todas as três variáveis de todos os dez examinadores passaram nos testes. Estes resultados evidenciam que as medianas das amostras de cada um dos examinadores podem ser consideradas iguais a mediana da amostra geral da divisão. Vale ressaltar ainda a maioria de valores elevados de p-valor, ou seja, os resultados passaram muito bem nos testes, sendo que os examinadores que apresentaram menores valores foram o 4 e 5, algo coerente com os resultados e análises da tabela 53.

5.4 TESTES DE REDISTRIBUIÇÃO – OUTRAS ÁREAS

Para a composição das amostras de redistribuição das demais áreas avaliadas, foram selecionados dez pedidos de patente de cinco examinadores de cada área, de modo que cada amostra tivesse pelo menos 50 pedidos de patente. Destaque-se ainda que, assim como no caso da amostra de redistribuição da área de eletricidade, como nestes casos o objetivo era obter uma amostra para aplicação do modelo validado com a amostra padrão com tempo (nossa referência), foram obtidas todas as variáveis de interesse de primeiros exames já publicados ao longo do segundo semestre de 2020, sendo que todos os pedidos da amostra também são do tipo 6.21, ou seja, com aproveitamento de buscas de escritórios internacionais.

Destaque-se que, de maneira análoga ao executado para o teste de redistribuição da área de eletricidade, as simulações aqui realizadas foram as do caso 3 Var (2). Os testes de Bartlett e KMO também foram aplicados para cada um dos casos. Quanto ao Teste de Bartlett, com exceção dos casos com somente 10 amostras, todos os demais foram aprovados. Além disso, todas as amostras foram aprovadas no teste KMO, desde 10 até 50 pedidos.

5.4.1 Amostra de Redistribuição – Área de Eletrônica

A Tabela 55 mostra, para o conjunto total de 50 pedidos, um quadro geral das principais variáveis obtidas e os resultados dos cálculos de estatísticas básicas tais como: média, desvio padrão, mediana, moda, mínimo, máximo e coeficiente de variação.

Tabela 55 – Estatísticas Básicas da Amostra de Redistribuição

	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Reivindicações Independentes	Reivindicações Dependentes	Total de Subclasses	Ano de Depósito	Inventores	Prioridades
Média	37,04	5,24	11,04	2,86	11,28	1,70	11,14	2,74	1,24
Desvio Padrão	30,96	2,70	11,47	1,39	5,90	0,99	1,05	1,87	1,24
Mediana	26,50	4,50	7,50	2,50	11,50	1,00	11,00	2,00	1,00
Moda	13	4	8	2	13	1	11	1	1
Mínimo	6	2	1	1	0	1	10	1	0
Máximo	170	14	62	7	30	5	13	8	9
Coeficiente de Variação	83,58%	51,52%	103,89%	48,43%	52,34%	58,52%	9,43%	68,30%	99,85%

Fonte: Elaboração Própria.

Da análise da Tabela 55, verifica-se que, para quase todas as variáveis do pedido a serem consideradas, os desvios padrão (e variância, pois é o quadrado do desvio) e coeficientes de variação (relação entre o desvio padrão e a média) são altos. A exceção é o ano de depósito que apresenta coeficiente de variação bem baixo (9,43%) e moda igual a 11, evidenciando que a distribuição original para os examinadores ficou concentrada em poucos anos de depósito, principalmente no ano de 2011. Adicionalmente, modas, médias e medianas das variáveis são distintas entre si, sendo que os valores mínimos e máximos são bastante diferentes. Tais resultados sugerem que alta amplitude e boa dispersão são características intrínsecas dos pedidos analisados, ou seja, há uma variabilidade razoável entre eles mesmo considerando pedidos de mesmo tipo (6.21 com aproveitamento de buscas), sendo que, conforme desejado, o perfil dos dados é parecido com o da amostra padrão com tempo e com o da amostra de redistribuição da área de eletricidade.

5.4.1.1 Execução da Análise de Componentes Principais: Caso 3 Var (2)

A Tabela 56 mostra os principais resultados da execução da ACP para os dados da área de eletrônica.

Tabela 56 – Autovalores e Contribuições de Variância das Componentes Principais

Componente Principal	Autovalores	Variância	Variância Acumulada
Y1	1,86	62%	62%
Y2	0,82	27%	89%
Y3	0,32	8%	100%

Fonte: Elaboração Própria.

Da análise da Tabela 56, verifica-se que no caso de utilização do critério considerando autovalores maiores do que um, ou mesmo o critério de variância acumulada (acima de 60%), selecionaríamos somente a primeira componente para a determinação do IGC, ou seja, $IGC_{\lambda>1} = IGC_{Y1}$.

As Figuras 47 e 48 mostram, respectivamente, os autovalores e as variâncias acumuladas para cada um dos casos 3 var (2), com amostras de 10 até 50 pedidos.

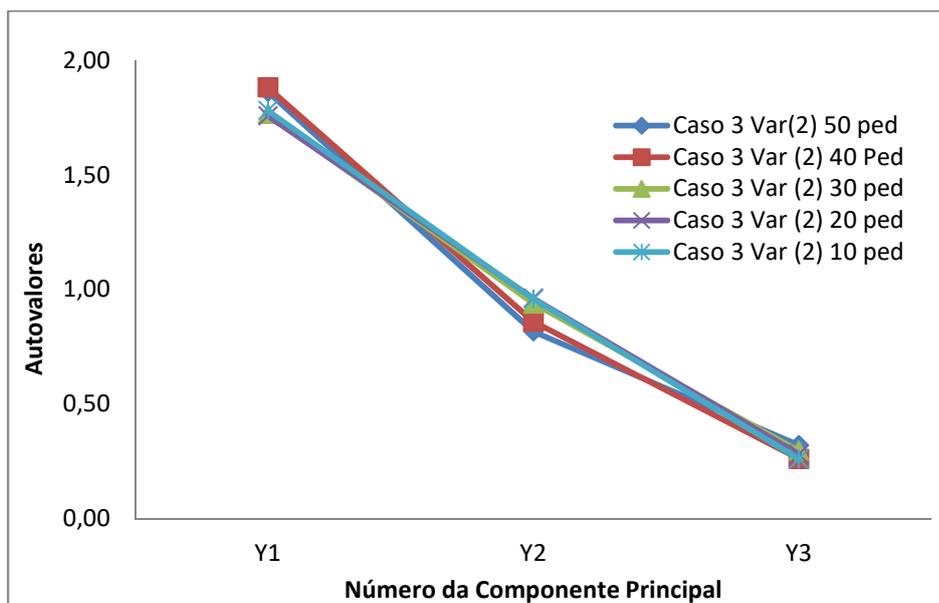


Figura 47 – Caso 3 Var (2): autovalores para amostras de 10 até 50 pedidos

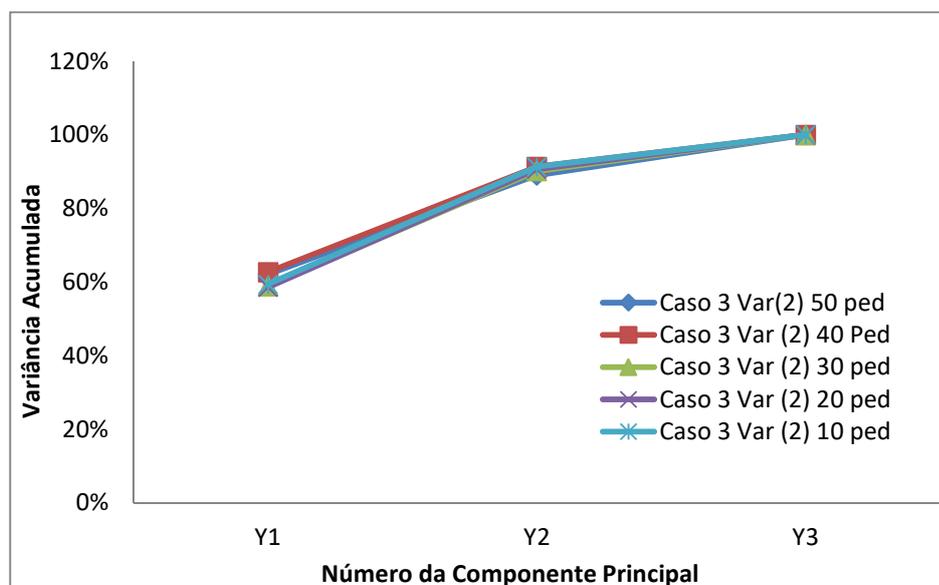


Figura 48 – Caso 3 Var (2): variância acumulada para amostras de 10 até 50 pedidos

Da análise da Figura 47, verifica-se que todos os casos, desde amostras de 10 até 50 pedidos, são praticamente coincidentes, ou seja, os autovalores das componentes apresentam uma excelente estabilidade. O mesmo fenômeno pode ser observado ao analisar as variâncias acumuladas das amostras conforme ilustrado na Figura 48.

Estes resultados reforçam que, para os casos com três variáveis que se pretenda executar, já é possível utilizar uma amostra de 10 pedidos de patente, sendo preferível a partir de 20 pedidos, de modo que se consiga um melhor rendimento do método de ACP.

A Tabela 57 mostra os escores das duas primeiras componentes principais para cada uma das variáveis consideradas do caso 3 var (2) com 50 pedidos.

Tabela 57 – Escores das Componentes Principais de cada variável

Variáveis	Coeficiente de Ponderação	
	Y1	Y2
Total de Páginas	0,4382	-0,8789
Total de Reivindicações	0,6136	0,4458
Páginas de Reivindicações	0,6569	0,1694

Fonte: Elaboração Própria.

Na Tabela 57 foram hachurados na cor cinza os valores considerados significativos dos fatores para cada uma das variáveis (próximos ou acima de 0,4). Em uma primeira análise, nota-se que a componente Y1 é bastante destacada das demais, sendo responsável por explicar 62% de toda a variância. Neste caso, ambas as componentes Y1 e Y2 estão associadas às principais variáveis diretamente relacionadas ao volume de dados (total de páginas, páginas de reivindicações, além de número de reivindicações). Este resultado é bem semelhante aos demais casos 3 var (2) já executados.

5.4.1.2 Enquadramento dos Pedidos de Patente da Amostra de Redistribuição por IGC_{Y1}

A Figura 49 mostra os IGC_{Y1} paronizados (eixo vertical) para todos os cinquenta pedidos (eixo horizontal) da amostra. As retas horizontais facilitam a verificação visual dos enquadramentos. Já a Figura 50 mostra a quantidade de enquadramentos por IGC_{Y1} nas classes.

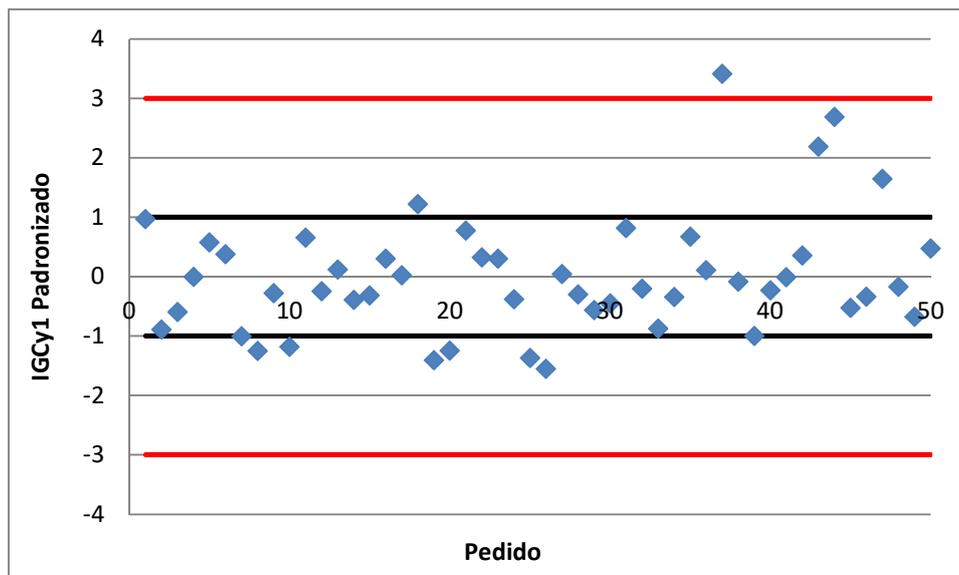


Figura 49 – IGC_{Y1} padronizado dos pedidos

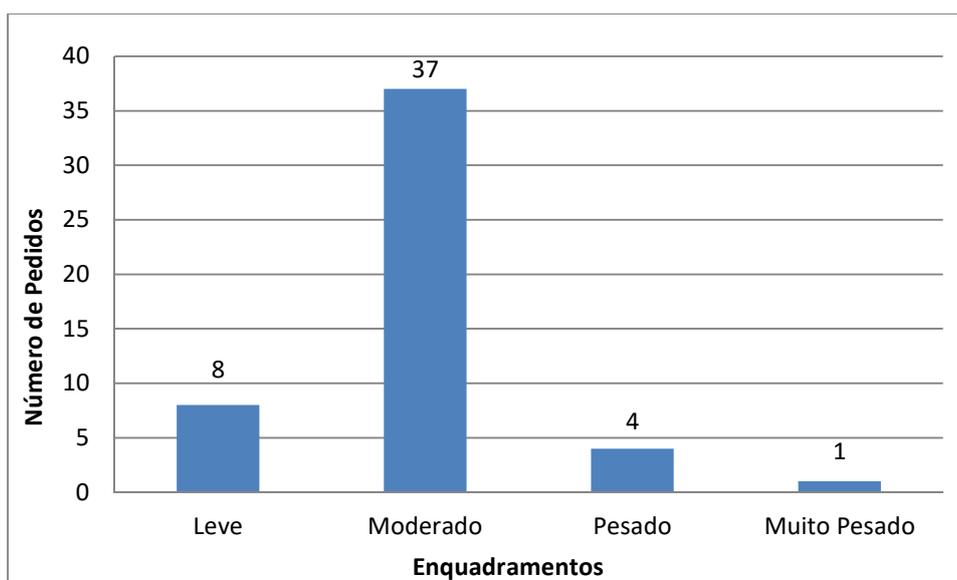


Figura 50 – Enquadramentos por IGC_{Y1} dos pedidos da amostra

Da análise das Figuras 49 e 50, verifica-se que nenhum dos pedidos de patente da amostra foi enquadrado como muito leve. A maioria dos pedidos foi enquadrada como moderado (74%), em segundo lugar como leve (16%), seguidos pelos pesados (8%) e, por fim, os muito pesados, que foram apenas um (2%). Vale ressaltar novamente que este resultado se mostra coerente com o dos testes anteriores e, assim como no caso da amostra padrão com tempo, neste caso a amostra de dados de IGC_{Y1} também apresentou distribuição normal.

As Tabelas 58 a 60 detalham os dados dos pedidos enquadrados como muito pesados, pesados e leves com relação ao IGC_{Y1} .

Tabela 58 – Pedidos de patente enquadrados como muito pesados por IGC_{Y1}

Pedidos Muito Pesados no Enquadramento por IGC_{Y1}							
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
37	130	11	43	184	4	30	34

Fonte: Elaboração Própria.

Tabela 59 – Pedidos de patente enquadrados como pesados por IGC_{Y1}

Pedidos Pesados no Enquadramento por IGC_{Y1}							
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
18	20	7	7	34	3	26	29
43	102	12	43	157	5	12	17
44	63	14	14	91	5	21	26
47	37	12	7	56	7	13	20

Fonte: Elaboração Própria.

Tabela 60 – Pedidos de patente enquadrados como leves por IGC_{Y1}

Pedidos Leves no Enquadramento por IGC_{Y1}							
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
7	13	3	5	21	1	8	9
8	14	2	5	21	1	7	8
10	26	4	9	39	2	0	2
19	13	2	4	19	2	4	6
20	8	2	2	12	1	8	9
25	29	2	12	43	1	3	4
26	6	2	1	9	2	3	5
39	13	2	2	17	2	10	12

Fonte: Elaboração Própria.

Como pode ser observado das Tabelas 58 a 60, conforme esperado os enquadramentos pelo IGC_{Y1} mostram um padrão específico quanto às três variáveis do caso simulado: número de páginas de reivindicações, número total de páginas e número total de reivindicações.

De uma análise específica da Tabela 58, verifica-se que o pedido 37 apresentou número de páginas de reivindicações (11), número total de reivindicações (34) e número total de páginas (184) muito acima das respectivas médias (5,24, 14,14 e 53,32). Consequentemente, o pedido foi enquadrado como muito pesado.

Considerações análogas às anteriores podem ser feitas para os pedidos enquadrados como pesados, conforme ilustra a Tabela 59. Com exceção do número total de páginas do

pedido 18, que apresenta valor abaixo da média geral, todas as demais variáveis dos pedidos 18, 43, 44 e 47 apresentaram valores acima da média geral. Todavia, tais valores não foram tão acima da média (mais do que três desvios padrão) a ponto de gerar enquadramentos muito pesados e, sendo assim, o modelo sugeriu um enquadramento como pesado.

Por fim, da análise da Tabela 60 também pode-se observar um padrão para as três variáveis avaliadas. Todos os pedidos apresentaram as três variáveis com valores razoavelmente abaixo da média. Todavia, os valores não foram tão abaixo da média a ponto de gerar algum enquadramento muito leve e, sendo assim, o modelo os enquadrrou como leves. Especificamente com relação aos pedidos 20 e 26, verifica-se que eles são bem pequenos, compreendendo duas páginas de reivindicações, total de somente doze e nove páginas, e somente um total de nove e cinco reivindicações. Principalmente quanto ao pedido 26, ele poderia ser enquadrado como muito leve ao invés de leve, evidenciando que, para situações em que as três variáveis apresentem coeficiente de variação muito alto, o modelo pode apresentar limitações para enquadrar pedidos muito leves. Mais uma vez, destaque-se que, em situações de dúvidas para enquadramentos muito leves, é sugerida uma análise complementar do percentil 5 da amostra. Os percentis 5 para as variáveis número de páginas de reivindicações, número total de páginas e número total de reivindicações foram, respectivamente, iguais a 2, 17,9 e 4,45. Verifica-se que ambos os pedidos estão abaixo do limite do percentil 5 para o total de páginas, sendo que o pedido 26 também apresenta as outras duas variáveis bem perto desse limite. Consequentemente, embora não seja indiscutível, tais pedidos poderiam ser enquadrados como muito leves, principalmente o 26, e a análise de percentil 5 pode auxiliar na decisão.

5.4.1.3 Análise das Subclasses da CIP da Amostra de Redistribuição

5.4.1.3.1 ZAE Original dos Examinadores da Amostra de Redistribuição

A Tabela 61 a seguir mostra a configuração das zonas de atuação de cada um dos 5 examinadores por subclasses da CIP para a distribuição original. Destaca-se que a área cinza corresponde às subclasses das Zonas de Atuação Específicas (ZAE) dos examinadores. Ressalte-se que, como neste caso a amostra é bem menor que a amostra de teste inicial, as ZAE foram determinadas considerando subclasses de pedidos de patente com ocorrências a partir de 20% do total de exames realizados por cada um dos examinadores avaliados.

Tabela 61 – Zonas de Atuação dos Examinadores por Subclasses da CIP

ZAE original da Amostra de Redistribuição									
1		2		3		4		5	
Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%
H04R	40,00%	G06F	70,00%	G06T	40,00%	G06F	50,00%	G10L	30,00%
G06Q	10,00%	H03F	10,00%	G06F	40,00%	G06Q	40,00%	G06F	30,00%
F42B	10,00%	G06Q	10,00%	G06Q	10,00%	HC3M	10,00%	G06Q	30,00%
G06K	10,00%	B60K	10,00%	G06K	10,00%			G06G	10,00%
H03L	10,00%								
G06F	10,00%								
H03B	10,00%								

Fonte: Elaboração Própria

Da análise da Tabela 61, verifica-se que as configurações de distribuição nas ZAE de cada um dos examinadores são bem variáveis. Enquanto os examinadores 4 e 5 apresentam nove dos seus dez pedidos distribuídos dentro das ZAE de referência, o examinador 1 apresenta quatro, o examinador 2 apresenta sete e o examinador 3 apresenta oito. Tal fato mostra que, mesmo com uma tentativa de distribuição de carga atual utilizando como base a especialidade/áreas de conhecimento dos examinadores, caso a amostra não contenha pedidos suficientes, ou mesmo essa distribuição não esteja sistematizada/automatizada de forma a auxiliar os gestores, podem ocorrer distorções indesejáveis. Por outro lado nota-se uma clara tentativa de distribuição tendo como referência as áreas de atuação dos examinadores.

5.4.1.3.2 Configuração Original dos Pedidos

A Tabela 62 mostra as medianas das variáveis dos pedidos com os dados separados por examinador. Nesta tabela foram hachurados: na cor cinza claro as variáveis dos pedidos de examinadores com mediana abaixo da mediana geral dos dados; e na cor cinza escuro as variáveis dos pedidos dos examinadores com mediana acima da mediana geral dos dados.

Tabela 62 – Medianas da Amostra de Redistribuição separadas por examinador

Medianas							
Examinador	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
1	24	4	5,5	36,5	2	10,5	12,5
2	20	4,5	6	31	3	11	15
3	29	4	10	42,5	2	11,5	13,5
4	29,5	4,5	8,5	41,5	3	11	13
5	44,5	6,5	11	64,5	3	12	14,5
Divisão	26,5	4,5	7,5	39,0	2,5	11,5	14,0

Fonte: Elaboração Própria.

Da análise da Tabela 62, verifica-se que praticamente todos os examinadores apresentam, para as três variáveis em análise, medianas tanto abaixo quanto acima da mediana geral dos dados da divisão. Nota-se, ainda, que o examinador 5 apresenta medianas do número de páginas de reivindicações e total de páginas muito acima das respectivas medianas gerais da divisão, sugerindo que pode estar com carga um pouco pesada em relação aos demais. Já o examinador 1 apresenta medianas de todas as três variáveis um pouco abaixo das medianas da divisão, sugerindo que pode estar com carga ligeiramente leve em relação aos demais. Em suma, *a priori* o balanceamento parece estar bom, com alguns desvios a tentar corrigir ou minimizar com uma nova redistribuição.

5.4.1.3.3 Redistribuição dos Pedidos

5.4.1.3.3.1 Nova Configuração das ZAE dos Examinadores

A Tabela 63 mostra a nova configuração das ZAE dos examinadores após ser realizada uma distribuição com a nova lógica desenvolvida.

Tabela 63 – ZAE dos examinadores após a nova distribuição

Novas ZAE da Amostra de Redistribuição - IGC _{v1}									
1		2		3		4		5	
Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%
H04R	30,00%	G06F	80,00%	G06T	40,00%	G06Q	80,00%	G06F	50,00%
G06F	20,00%	H03B	10,00%	G06F	30,00%	G06F	20,00%	G10L	30,00%
B60K	10,00%	H03F	10,00%	G06K	10,00%			G06Q	10,00%
H03L	10,00%			F42B	10,00%			H04R	10,00%
H03M	10,00%			G06K	10,00%				
G06G	10,00%								
G06Q	10,00%								

Fonte: Elaboração Própria

A Tabela 64 mostra uma comparação das porcentagens de pedidos dentro das ZAE em relação ao total de pedidos distribuídos para cada examinador nas distribuições nova e original.

Tabela 64 – Porcentagem de pedidos dentro das ZAE antes e após a nova distribuição

Examinador	1	2	3	4	5
Nova ZAE	30,00%	80,00%	70,00%	100,00%	90,00%
ZAE Original	40,00%	70,00%	80,00%	90,00%	90,00%

Fonte: Elaboração Própria

Da análise das Tabelas 63 e 64, verifica-se que ficou mantido o mesmo equilíbrio na concentração de exame dentro das ZAE de cada um dos examinadores. Destaque-se que somente o examinador 1 se manteve com uma pequena concentração (30% ou 40%) de exame dentro da sua ZAE, o que talvez possa ser explicado pelo fato de este ser um examinador mais versátil da divisão, que não tem uma ZAE enxuta e bem definida. Sendo assim, os resultados evidenciam que o equilíbrio foi mantido, porém agora com uma execução automática da distribuição, onde haverá uma tendência de melhora com amostras maiores e ZAE definidas com maior extensão e exatidão (o limite de subclasses de ZAE por examinador atualmente no programa é de três), onde se tenha mais margem para distribuição por subclasses da CIP.

5.4.1.3.3.2 Nova Configuração dos Pedidos Redistribuídos

A Tabela 65 mostra as medianas das variáveis dos pedidos redistribuídos com os dados separados por examinador. Nesta tabela foram hachurados: na cor cinza claro as variáveis dos pedidos de examinadores com mediana abaixo da mediana geral dos dados; e na cor cinza escuro as variáveis dos pedidos dos examinadores com mediana acima da mediana geral dos dados.

Tabela 65 – Medianas da Amostra de Redistribuição separadas por examinador

medianas							
Examinador	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
1	33	4	7	44	2	9	11
2	25	5	6,5	37,5	2	10	12
3	21,5	4	6	31,5	2,5	13	16
4	30	5,5	8	41,5	3,5	12,5	17
5	21	4,5	8,5	35	3	11	15,5
Geral	26,5	4,5	7,5	39,0	2,5	11,5	14,0

Fonte: Elaboração Própria.

Da análise da Tabela 65, verifica-se que ainda há algumas variáveis de pedidos dos examinadores com mediana tanto abaixo quanto acima da mediana geral dos dados da divisão. Todavia, nota-se que agora estes desvios foram reduzidos, principalmente com relação às variáveis número total de páginas e número de páginas de reivindicações. Nenhum examinador apresentou medianas das três variáveis abaixo da mediana geral, sendo que o examinador 4 foi o único que apresentou medianas das três variáveis acima da mediana geral, sendo uma delas próxima da mediana geral, sugerindo que possa estar levemente sobrecarregado em relação aos demais. Sendo assim, *a priori* não foi identificado um

examinador com tendências claras de carga leve ou sobrecarga com relação aos demais. Estes resultados evidenciam que, além de manter o equilíbrio das ZAE, a nova redistribuição teve um efeito positivo no ajuste fino das medianas das amostras dos examinadores, algo que será quantificado por meio do cálculo do IBD e avaliado por testes de hipótese.

5.4.1.3.3.3 Índice de Balanceamento da Distribuição (IBD) antes e após a Redistribuição

Foram calculados os índices IBD da distribuição original da amostra e da sua redistribuição. Para realização dos cálculos foram utilizadas as medianas constantes da tabela 62 (distribuição original) e da tabela 65 (redistribuição).

Por meio da equação (16), foram então obtidos IBD igual a 0,83 para o caso original e IBD igual a 0,88 para a redistribuição, ou seja, houve um aumento de 6,02 % do IBD com nova distribuição, demonstrando que as medianas dos pedidos dos examinadores após a redistribuição se aproximam mais da mediana geral da divisão. Tal fato corrobora que com a nova distribuição temos uma tendência de maior equilíbrio quanto ao volume de dados e tempo/esforço dos pedidos distribuídos aos examinadores.

5.4.1.3.3.4 Testes de Hipótese

A Tabela 66 mostra os resultados de p-valor dos testes de Mann-Whitney realizados para cada uma das três variáveis e cada um dos cinco examinadores.

Tabela 66 – Resultados de p-valor dos Testes de Mann-Whitney

Examinador	p-valor dos testes de Mann-Whitney		
	Páginas de Reivindicações	Total de Reivindicações	Total de Páginas
1	0,226	0,313	0,503
2	0,749	0,165	0,522
3	0,952	0,342	0,390
4	0,194	0,208	0,430
5	0,873	0,857	0,976

Fonte: Elaboração Própria.

Da análise da Tabela 66, verifica-se que as amostras de todas as três variáveis de todos os cinco examinadores passaram nos testes. Estes resultados evidenciam que as medianas das amostras de cada um dos examinadores podem ser consideradas iguais a mediana da amostra geral da divisão. Vale ressaltar ainda a maioria de valores elevados de p-valor, ou seja, os resultados passaram muito bem nos testes, sendo que o examinador que apresentou menores valores foi o 4, algo coerente com os resultados e análises da tabela 65.

5.4.2 Amostra de Redistribuição – Área de Química (Polímeros)

A Tabela 67 mostra, para o conjunto total de 50 pedidos, um quadro geral das principais variáveis obtidas e os resultados dos cálculos de estatísticas básicas tais como: média, desvio padrão, mediana, moda, mínimo, máximo e coeficiente de variação.

Tabela 67 – Estatísticas Básicas da Amostra de Redistribuição

	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Reivindicações Independentes	Reivindicações Dependentes	Total de Subclasses	Ano de Depósito	Inventores	Prioridades
Média	31,46	4,18	3,96	2,98	11,92	1,92	11,90	3,64	1,12
Desvio Padrão	22,54	2,63	6,11	1,72	5,81	0,90	1,25	2,27	0,52
Mediana	25,00	3,00	2,00	3,00	11,00	2,00	12,00	3,00	1,00
Moda	22	3	0	3	7	2	12	3	1
Mínimo	6	1	0	1	2	1	9	1	0
Máximo	108	15	36	10	26	5	16	12	3
Coeficiente de Variação	71,65%	62,96%	154,32%	57,72%	48,71%	46,87%	10,50%	62,24%	46,48%

Fonte: Elaboração Própria.

Da análise da Tabela 67, verifica-se que, para quase todas as variáveis do pedido a serem consideradas, os desvios padrão (e variância, pois é o quadrado do desvio) e coeficientes de variação (relação entre o desvio padrão e a média) são altos. A exceção é o ano de depósito que apresenta coeficiente de variação baixo (10,50%) e moda igual a 12, evidenciando que a distribuição original para os examinadores ficou concentrada em poucos anos de depósito, principalmente no ano de 2012. Adicionalmente, modas, médias e medianas da maioria das variáveis são distintas entre si, sendo que os valores mínimos e máximos são bastante diferentes. Tais resultados sugerem que alta amplitude e boa dispersão são características intrínsecas dos pedidos analisados, ou seja, há uma variabilidade razoável entre eles mesmo considerando pedidos de mesmo tipo (6.21 com aproveitamento de buscas), sendo que, conforme desejado, o perfil dos dados é parecido com o da amostra padrão com tempo e com o das amostras de redistribuição das áreas de eletricidade e eletrônica.

5.4.2.1 Execução da Análise de Componentes Principais: Caso 3 Var (2)

A Tabela 68 mostra os principais resultados da execução da ACP para os dados da área de química (polímeros).

Tabela 68 – Autovalores e Contribuições de Variância das Componentes Principais

Componente Principal	Autovalores	Variância	Variância Acumulada
Y1	2,15	72%	72%
Y2	0,60	20%	92%
Y3	0,25	8%	100%

Fonte: Elaboração Própria.

Da análise da Tabela 68, verifica-se que no caso de utilização do critério considerando autovalores maiores do que um, ou mesmo o critério de variância acumulada (acima de 60%), selecionaríamos somente a primeira componente para a determinação do IGC, ou seja, $IGC_{\lambda>1} = IGC_{Y1}$. As Figuras 51 e 52 mostram, respectivamente, os autovalores e as variâncias acumuladas para cada um dos casos 3 var (2), com amostras de 10 até 50 pedidos.

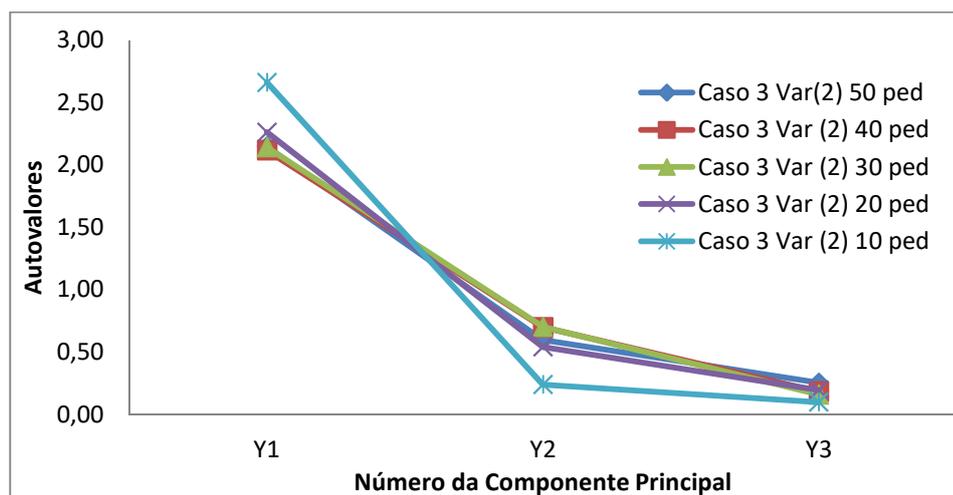


Figura 51 – Caso 3 Var (2): autovalores para amostras de 10 até 50 pedidos

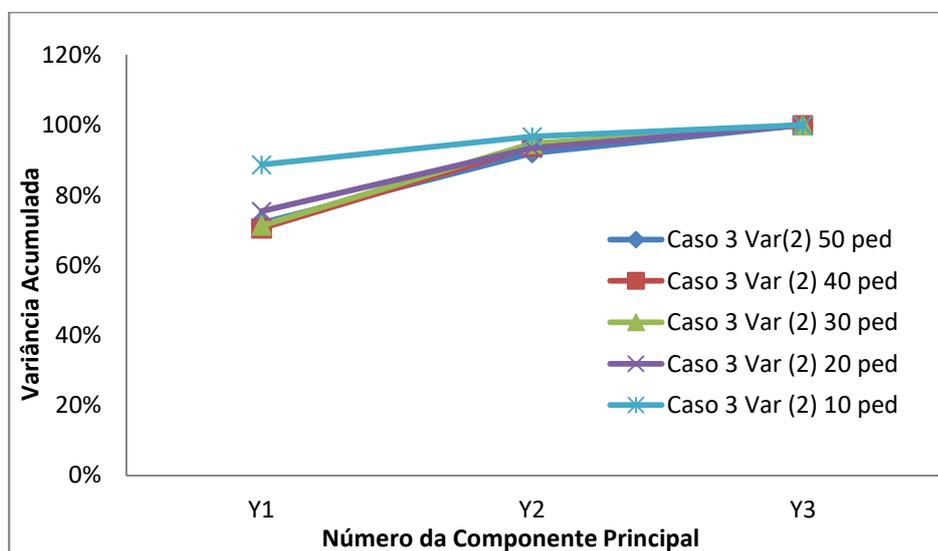


Figura 52 – Caso 3 Var (2): variância acumulada para amostras de 10 até 50 pedidos

Da análise da Figura 51, verifica-se que os casos com amostras de 20 até 50 pedidos são praticamente coincidentes, ou seja, os autovalores das componentes apresentam uma excelente estabilidade. O mesmo fenômeno pode ser observado ao analisar as variâncias acumuladas das amostras conforme ilustrado na Figura 52. Estes resultados reforçam que,

para os casos com três variáveis que se pretenda executar, é recomendável utilizar uma amostra com pelo menos 20 pedidos de patente.

A Tabela 69 mostra os escores das duas primeiras componentes principais para cada uma das variáveis consideradas do caso 3 var (2) com 50 pedidos.

Tabela 69 – Escores das Componentes Principais de cada variável

Variáveis	Coeficiente de Ponderação	
	Y1	Y2
Total de Páginas	0,5312	-0,7849
Total de Reivindicações	0,5713	0,6099
Páginas de Reivindicações	0,6257	0,1093

Fonte: Elaboração Própria.

Na Tabela 69 foram hachurados na cor cinza os valores considerados significativos dos fatores para cada uma das variáveis (próximos ou acima de 0,4). Em uma primeira análise, nota-se que a componente Y1 é bastante destacada das demais, sendo responsável por explicar 72% de toda a variância. Neste caso, ambas as componentes Y1 e Y2 estão associadas às principais variáveis diretamente relacionadas ao volume de dados (total de páginas, páginas de reivindicações, além de número de reivindicações). Este resultado é bem semelhante aos demais casos 3 var (2) já executados.

5.4.2.2 Enquadramento dos Pedidos de Patente da Amostra Final de Redistribuição por IGC_{Y1}

A Figura 53 mostra os IGC_{Y1} padronizados (eixo vertical) para todos os cinquenta pedidos (eixo horizontal) da amostra. As retas horizontais facilitam a verificação visual dos enquadramentos. Já a Figura 54 mostra a quantidade de enquadramentos por IGC_{Y1} nas classes.

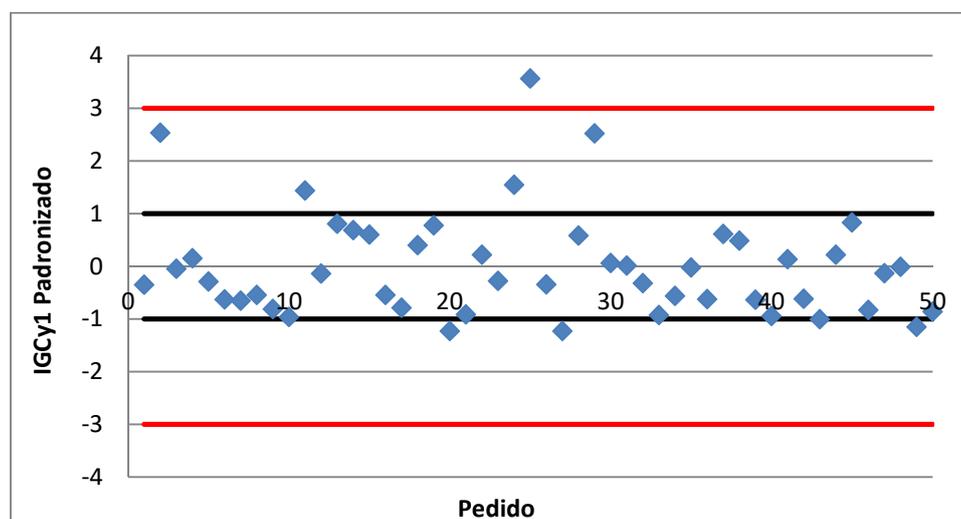


Figura 53 – IGC_{Y1} padronizado dos pedidos

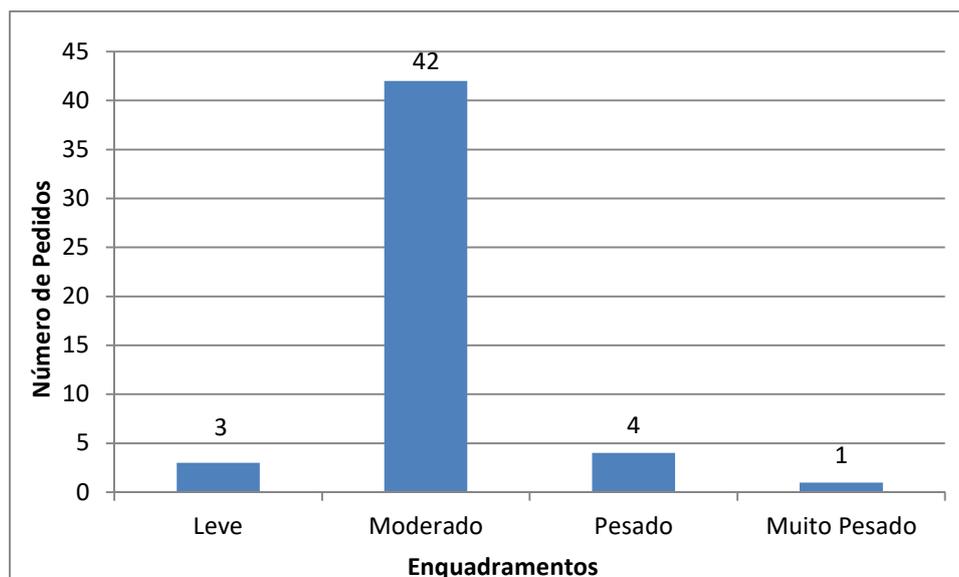


Figura 54 – Enquadramentos por IGC_{Y1} dos pedidos da amostra

Da análise das Figuras 53 e 54, verifica-se que nenhum dos pedidos de patente da amostra foi enquadrado como muito leve. A maioria dos pedidos foi enquadrada como moderado (84%), em segundo lugar como pesado (8%), seguidos pelos leves (6%) e, por fim, os muito pesados, que foram apenas um (2%). Vale ressaltar novamente que este resultado se mostra coerente com o dos testes anteriores e, assim como no caso da amostra padrão com tempo, neste caso a amostra de dados de IGC_{Y1} também apresentou distribuição normal.

As Tabelas 70 a 72 detalham os dados dos pedidos enquadrados como muito pesados, pesados e leves com relação ao IGC_{Y1}.

Tabela 70 – Pedido de patente enquadrado como muito pesado por IGC_{Y1}

Pedidos Muito Pesados no Enquadramento por IGC _{Y1}							
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
25	108	15	9	132	2	17	19

Fonte: Elaboração Própria.

Tabela 71 – Pedidos de patente enquadrados como pesados por IGC_{Y1}

Pedidos Pesados no Enquadramento por IGC _{Y1}							
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
2	76	10	8	94	6	21	27
11	25	7	12	44	5	26	31
24	72	6	1	79	6	20	26
29	99	12	2	113	3	12	15

Fonte: Elaboração Própria.

Tabela 72 – Pedidos de patente enquadrados como leves por IGC_{Y1}

Pedidos Leves no Enquadramento por IGC _{Y1}							
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
20	6	2	2	10	1	7	8
27	8	2	0	10	1	7	8
49	16	2	4	22	2	4	6

Fonte: Elaboração Própria.

Como pode ser observado das Tabelas 70 a 72, conforme esperado os enquadramentos pelo IGC_{Y1} mostram um padrão específico quanto às três variáveis do caso simulado: número de páginas de reivindicações, número total de páginas e número total de reivindicações.

De uma análise específica da Tabela 70, verifica-se que o pedido 25 apresentou número de páginas de reivindicações (15) e número total de páginas (132) muito acima das respectivas médias (4,18 e 39,6). Adicionalmente, o seu número total de reivindicações (19) também é acima da média (14,9). Consequentemente, o pedido foi enquadrado como muito pesado.

Considerações análogas às anteriores podem ser feitas para os pedidos enquadrados como pesados, conforme ilustra a Tabela 71. Cabe destacar que todas as três variáveis relevantes dos pedidos 2, 11, 24 e 29 apresentaram valores acima da média geral, sendo pelo menos duas delas razoavelmente acima da média. Todavia, tais valores não foram tão acima da média (mais do que três desvios padrão) a ponto de gerar enquadramentos muito pesados e, sendo assim, o modelo sugeriu um enquadramento como pesado.

Por fim, da análise da Tabela 72 também pode ser observado um padrão para as três variáveis avaliadas. Todos os pedidos apresentaram as três variáveis com valores razoavelmente abaixo da média. Todavia, os valores não foram tão abaixo da média a ponto de gerar algum enquadramento muito leve e, sendo assim, o modelo os enquadrado como leves. Especificamente com relação aos pedidos 20 e 27, verifica-se que eles são bem pequenos, compreendendo duas páginas de reivindicações, total de dez páginas e oito reivindicações. Ressalte-se que os percentis 5 para as variáveis número de páginas de reivindicações, número total de páginas e número total de reivindicações foram, respectivamente, iguais a 2, 12,9 e 6,45. Verifica-se que ambos os pedidos estão no limite do percentil 5 quanto ao número de páginas de reivindicações, abaixo do limite para o total de páginas, mas acima do limite para o total de reivindicações. Consequentemente, embora não seja indiscutível, tais pedidos poderiam ser enquadrados como muito leves, e a análise de percentil 5 pode auxiliar no enquadramento.

5.4.2.3 Análise das Subclasses da CIP da Amostra de Redistribuição

5.4.2.3.1 ZAE Original dos Examinadores da Amostra de Redistribuição

A Tabela 73 a seguir mostra a configuração das zonas de atuação de cada um dos 5 examinadores por subclasses da CIP para a distribuição original. Destaca-se que a área cinza corresponde às subclasses das Zonas de Atuação Específicas (ZAE) dos examinadores. Ressalte-se que, assim como no caso da amostra da área de eletrônica, as ZAE foram determinadas considerando subclasses de pedidos de patente com ocorrências a partir de 20% do total de exames realizados por cada um dos examinadores avaliados.

Tabela 73 – Zonas de Atuação dos Examinadores por Subclasses da CIP

ZAE original da Amostra de Redistribuição									
1		2		3		4		5	
Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%
C08F	50,00%	B29C	50,00%	C08G	30,00%	B29C	60,00%	C08G	30,00%
C08G	20,00%	B32B	40,00%	C08L	20,00%	B29D	20,00%	C08L	20,00%
C08K	10,00%	B07B	10,00%	C08F	20,00%	B32B	20,00%	C08F	20,00%
B32B	10,00%			C08B	10,00%			C08J	10,00%
C08L	10,00%			B32B	10,00%			B29C	10,00%
				C08C	10,00%			B32B	10,00%

Fonte: Elaboração Própria

Da análise da Tabela 73, verifica-se que as configurações de distribuição nas ZAE de cada um dos examinadores são razoavelmente concentradas. Enquanto os examinadores 1, 3 e 5 apresentam sete dos seus dez pedidos distribuídos dentro das ZAE de referência, o examinador 2 apresenta nove e o examinador 4 apresenta os dez. Tal fato evidencia uma tentativa de distribuição de carga atual utilizando como base a especialidade/áreas de conhecimento dos examinadores, sendo que a amostra avaliada parece contribuir para tal, dado que apresentou variedade de pedidos com subclasses das ZAE para distribuição.

5.4.2.3.2 Configuração Original dos Pedidos

A Tabela 74 mostra as medianas das variáveis dos pedidos com os dados separados por examinador. Nesta tabela foram hachurados: na cor cinza claro as variáveis dos pedidos de examinadores com mediana abaixo da mediana geral dos dados; e na cor cinza escuro as variáveis dos pedidos dos examinadores com mediana acima da mediana geral dos dados.

Tabela 74 – Medianas da Amostra de Redistribuição separadas por examinador

Medianas							
Examinador	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
1	33	3	0	36,5	3,5	8	11,5
2	20,5	5	6,5	35,5	3	17,5	21
3	30	4,5	1	37,5	3	13,5	15
4	21,5	3	4	31	2	12	14,5
5	26,5	3	0,5	36	2,5	7,5	10,5
Divisão	25,0	3,0	2,0	35,0	3,0	11,0	14,5

Fonte: Elaboração Própria.

Da análise da Tabela 74, verifica-se que praticamente todos os examinadores apresentam, para pelo menos duas das três variáveis em análise, medianas tanto abaixo quanto acima da mediana geral dos dados da divisão. Nota-se ainda que os examinadores 2 e 3 apresentam as medianas das três variáveis acima das respectivas medianas gerais da divisão, sendo que o número de páginas de reivindicações bem acima, sugerindo que os mesmos podem estar com carga um pouco pesada em relação aos demais. Em suma, a *priori* o balanceamento parece estar bom, com alguns desvios a tentar corrigir ou minimizar com uma nova redistribuição.

5.4.2.3.3 Redistribuição dos Pedidos

5.4.2.3.3.1 Nova Configuração das ZAE dos Examinadores

A Tabela 75 mostra a nova configuração das ZAE dos examinadores após ser realizada uma distribuição com a nova lógica desenvolvida.

Tabela 75 – ZAE dos examinadores após a nova distribuição

Novas ZAE da Amostra de Redistribuição - IGC _{v1}									
1		2		3		4		5	
Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%
C08F	50,00%	B29C	50,00%	C08L	30,00%	B29C	50,00%	C08G	30,00%
C08G	30,00%	B32B	50,00%	C08G	20,00%	B29D	20,00%	C08L	20,00%
C08J	10,00%			C08F	20,00%	B32B	20,00%	C08F	20,00%
B32B	10,00%			B07B	10,00%	C08C	10,00%	B29C	20,00%
				C08B	10,00%			C08K	10,00%
				B32B	10,00%				

Fonte: Elaboração Própria

A Tabela 76 mostra uma comparação das porcentagens de pedidos dentro das ZAE em relação ao total de pedidos distribuídos para cada examinador nas distribuições nova e original.

Tabela 76 – Porcentagem de pedidos dentro das ZAE antes e após a nova distribuição

Examinador	1	2	3	4	5
Nova ZAE	80,00%	100,00%	70,00%	90,00%	70,00%
ZAE Original	70,00%	90,00%	70,00%	100,00%	70,00%

Fonte: Elaboração Própria

Da análise das Tabelas 75 e 76, verifica-se que o equilíbrio na concentração de exame dentro das ZAE de cada um dos examinadores que já era muito bom (vide tabela 73), foi ainda ligeiramente melhorado, dado que agora não temos mais três examinadores (1, 3 e 5) com a menor concentração de pedidos dentro da ZAE (70%), mas apenas dois examinadores (3 e 5).

5.4.2.3.3.2 Nova Configuração dos Pedidos Redistribuídos

A Tabela 77 mostra as medianas das variáveis dos pedidos redistribuídos com os dados separados por examinador. Nesta tabela foram hachurados: na cor cinza claro as variáveis dos pedidos de examinadores com mediana abaixo da mediana geral dos dados; e na cor cinza escuro as variáveis dos pedidos dos examinadores com mediana acima da mediana geral dos dados.

Tabela 77 – Medianas da Amostra de Redistribuição separadas por examinador

medianas							
Examinador	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
1	37	4	2	44,5	3	11,5	15
2	21,5	4	5	31	2	13,5	17,5
3	28,5	3	1	31	3	11	14,5
4	24,5	3,5	5,5	38	2,5	14,5	16,5
5	21	2,5	0	25,5	3	7	10,5
Geral	25,0	3,0	2,0	35,0	3,0	11,0	14,5

Fonte: Elaboração Própria.

Da análise da Tabela 77, verifica-se que ainda há algumas variáveis de pedidos dos examinadores com mediana tanto abaixo quanto acima da mediana geral dos dados da divisão. Todavia, nota-se que agora estes desvios foram reduzidos, principalmente com relação às variáveis número total de páginas e número de páginas de reivindicações. Embora

com diferenças pequenas para duas das três variáveis, o examinador 1 apresentou medianas das três variáveis acima da mediana geral, sugerindo que possa estar levemente sobrecarregado em relação aos demais. Já o examinador 5 foi o único que apresentou medianas das três variáveis abaixo da mediana geral, sendo uma delas (páginas de reivindicações) próxima da mediana geral, sugerindo que possa estar com carga um pouco leve em relação aos demais. Estes resultados evidenciam que, além de melhorar um pouco o equilíbrio das ZAE (vide tabelas 75 e 76), a nova redistribuição teve um efeito positivo no ajuste fino das medianas das amostras dos examinadores, algo que será quantificado por meio do cálculo do IBD e avaliado por testes de hipótese.

5.4.2.3.3.3 Índice de Balanceamento da Distribuição (IBD) antes e após a Redistribuição

Foram calculados os índices IBD da distribuição original da amostra e da sua redistribuição. Para realização dos cálculos foram utilizadas as medianas constantes da tabela 74 (distribuição original) e da tabela 77 (redistribuição).

Por meio da equação (16), foram então obtidos IBD igual a 0,81 para o caso original e IBD igual a 0,84 para a redistribuição, ou seja, houve um aumento de 3,7 % do IBD com nova distribuição, demonstrando que as medianas dos pedidos dos examinadores após a redistribuição se aproximam mais da mediana geral da divisão. Tal fato corrobora que com a nova distribuição temos uma tendência de maior equilíbrio quanto ao volume de dados e tempo/esforço dos pedidos distribuídos aos examinadores.

5.4.2.3.3.4 Testes de Hipótese

A Tabela 78 mostra os resultados de p-valor dos testes de Mann-Whitney realizados para cada uma das três variáveis e cada um dos cinco examinadores.

Tabela 78 – Resultados de p-valor dos Testes de Mann-Whitney

Examinador	p-valor dos testes de Mann-Whitney		
	Páginas de Reivindicações	Total de Reivindicações	Total de Páginas
1	0,187	0,897	0,056
2	0,327	0,238	0,441
3	0,928	0,944	0,976
4	0,881	0,435	0,920
5	0,018	0,056	0,327

Fonte: Elaboração Própria.

Da análise da Tabela 78, verifica-se que as amostras de todas as três variáveis de todos os cinco examinadores passaram nos testes. Estes resultados evidenciam que as medianas das amostras de cada um dos examinadores podem ser consideradas iguais a mediana da amostra geral da divisão. Vale ressaltar ainda a maioria de valores elevados de p-valor, ou seja, os resultados passaram muito bem nos testes, sendo que os examinadores que apresentaram menores valores foram o 1 e o 5, algo coerente com os resultados e análises da tabela 77.

5.4.3 Amostra de Redistribuição – Área de Biologia (Biofármacos)

A Tabela 79 mostra, para o conjunto total de 50 pedidos, um quadro geral das principais variáveis obtidas e os resultados dos cálculos de estatísticas básicas tais como: média, desvio padrão, mediana, moda, mínimo, máximo e coeficiente de variação.

Tabela 79 – Estatísticas Básicas da Amostra de Redistribuição

	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Reivindicações Independentes	Reivindicações Dependentes	Total de Subclasses	Ano de Depósito	Inventores	Prioridades
Média	57,84	4,08	10,50	3,64	12,54	2,04	10,22	3,84	1,18
Desvio Padrão	53,12	3,42	13,40	2,56	9,58	0,95	2,29	2,44	0,60
Mediana	40,50	3,00	6,00	3,00	10,50	2,00	11,00	3,00	1,00
Moda	22	2	0	2	4	2	11	2	1
Mínimo	1	1	0	1	0	1	4	1	0
Máximo	239	15	68	14	44	4	14	12	3
Coeficiente de Variação	91,84%	83,86%	127,66%	70,37%	76,39%	46,41%	22,39%	63,43%	50,47%

Fonte: Elaboração Própria.

Da análise da Tabela 79, verifica-se que, para quase todas as variáveis do pedido a serem consideradas, os desvios padrão (e variância, pois é o quadrado do desvio) e coeficientes de variação (relação entre o desvio padrão e a média) são altos. A exceção é o ano de depósito que apresenta coeficiente de variação moderado (22,39%) e moda igual a 11, evidenciando que a distribuição original para os examinadores ficou concentrada em poucos anos de depósito, principalmente no ano de 2011. Adicionalmente, modas, médias e medianas da maioria das variáveis são distintas entre si, sendo que os valores mínimos e máximos são bastante diferentes. Tais resultados sugerem que alta amplitude e boa dispersão são características intrínsecas dos pedidos analisados, ou seja, há uma variabilidade razoável entre eles mesmo considerando pedidos de mesmo tipo (6.21 com aproveitamento de buscas), sendo que, conforme desejado, o perfil dos dados é parecido com o da amostra padrão com tempo e com o das amostras de redistribuição das áreas de eletricidade, eletrônica e química.

5.4.3.1 Execução da Análise de Componentes Principais: Caso 3 Var (2)

A Tabela 80 mostra os principais resultados da execução da ACP para os dados da área de biologia (biofármacos).

Tabela 80 – Autovalores e Contribuições de Variância das Componentes Principais

Componente Principal	Autovalores	Variância	Variância Acumulada
Y1	2,03	68%	68%
Y2	0,74	25%	92%
Y3	0,23	8%	100%

Fonte: Elaboração Própria.

Da análise da Tabela 80, verifica-se que no caso de utilização do critério considerando autovalores maiores do que um, ou mesmo o critério de variância acumulada (acima de 60%), selecionaríamos somente a primeira componente para a determinação do IGC, ou seja, $IGC_{\lambda>1} = IGC_{Y1}$. As Figuras 55 e 56 mostram, respectivamente, os autovalores e as variâncias acumuladas para cada um dos casos 3 var (2), com amostras de 10 até 50 pedidos.

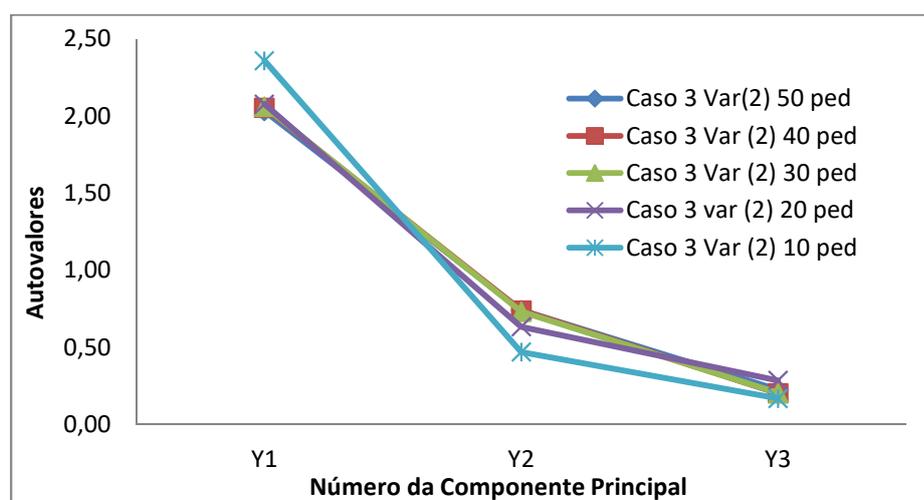


Figura 55 – Caso 3 Var (2): autovalores para amostras de 10 até 50 pedidos

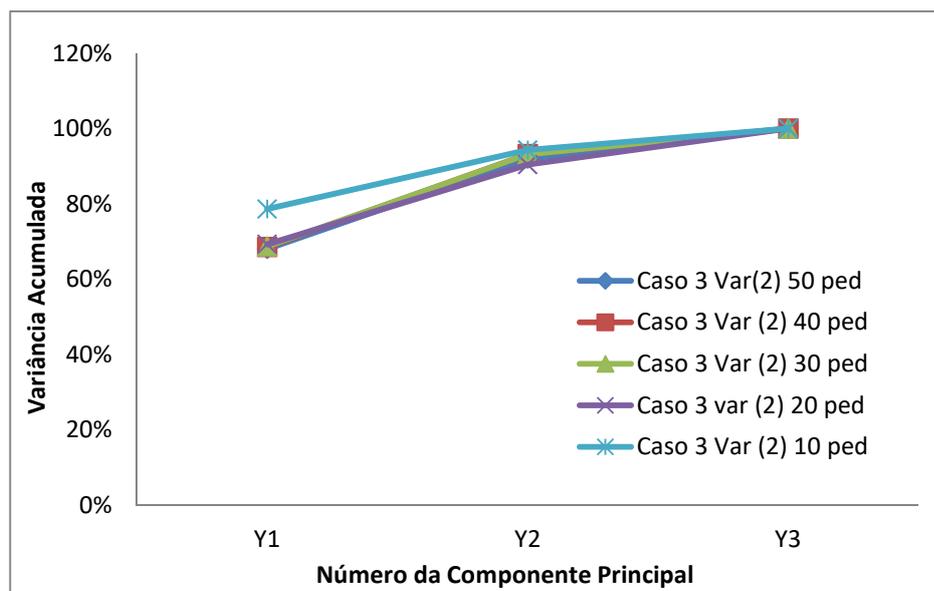


Figura 56 – Caso 3 Var (2): variância acumulada para amostras de 10 até 50 pedidos

Da análise da Figura 55, verifica-se que os casos com amostras de 20 até 50 pedidos são praticamente coincidentes, ou seja, os autovalores das componentes apresentam uma excelente estabilidade. O mesmo fenômeno pode ser observado ao analisar as variâncias acumuladas das amostras conforme ilustrado na Figura 56. Estes resultados reforçam que, para os casos com três variáveis que se pretenda executar, é recomendável utilizar uma amostra com pelo menos 20 pedidos de patente.

A Tabela 81 mostra os escores das duas primeiras componentes principais para cada uma das variáveis consideradas do caso 3 var (2) com 50 pedidos.

Tabela 81 – Escores das Componentes Principais de cada variável

Variáveis	Coeficiente de Ponderação	
	Y1	Y2
Total de Páginas	0,4597	-0,8784
Total de Reivindicações	0,6120	0,4256
Páginas de Reivindicações	0,6436	0,2176

Fonte: Elaboração Própria.

Na Tabela 81 foram hachurados na cor cinza os valores considerados significativos dos fatores para cada uma das variáveis (próximos ou acima de 0,4). Em uma primeira análise, nota-se que a componente Y1 é bastante destacada das demais, sendo responsável por explicar 68% de toda a variância. Neste caso, ambas as componentes Y1 e Y2 estão associadas às principais variáveis diretamente relacionadas ao volume de dados (total de páginas, páginas de reivindicações, além de número de reivindicações). Este resultado é bem semelhante aos demais casos 3 var (2) já executados.

5.4.3.2 Enquadramento dos Pedidos de Patente da Amostra de Redistribuição por IGC_{Y1}

A Figura 57 mostra os IGC_{Y1} padronizados (eixo vertical) para todos os cinquenta pedidos (eixo horizontal) da amostra. As retas horizontais facilitam a verificação visual dos enquadramentos. Já a Figura 58 mostra a quantidade de enquadramentos por IGC_{Y1} nas classes.

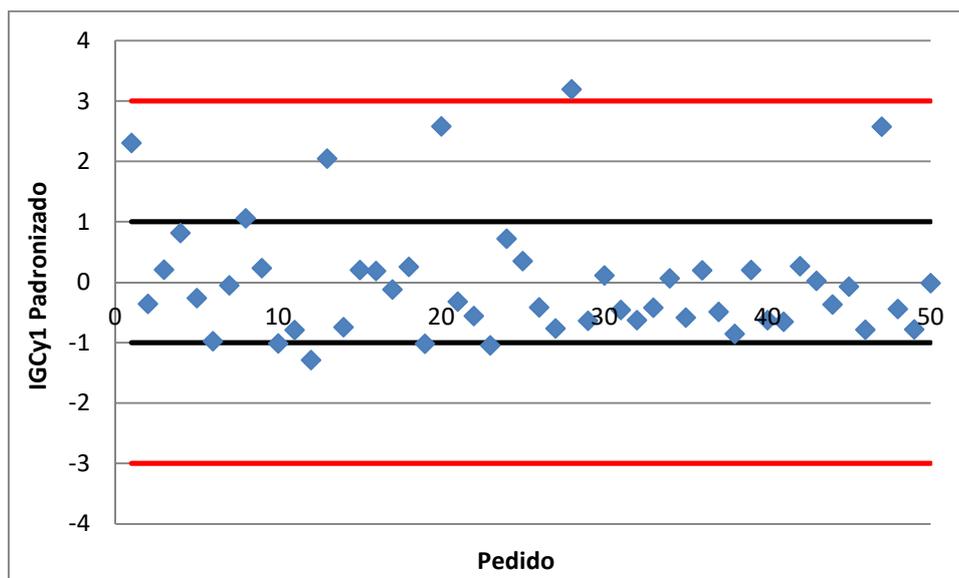


Figura 57 – IGC_{Y1} padronizado dos pedidos

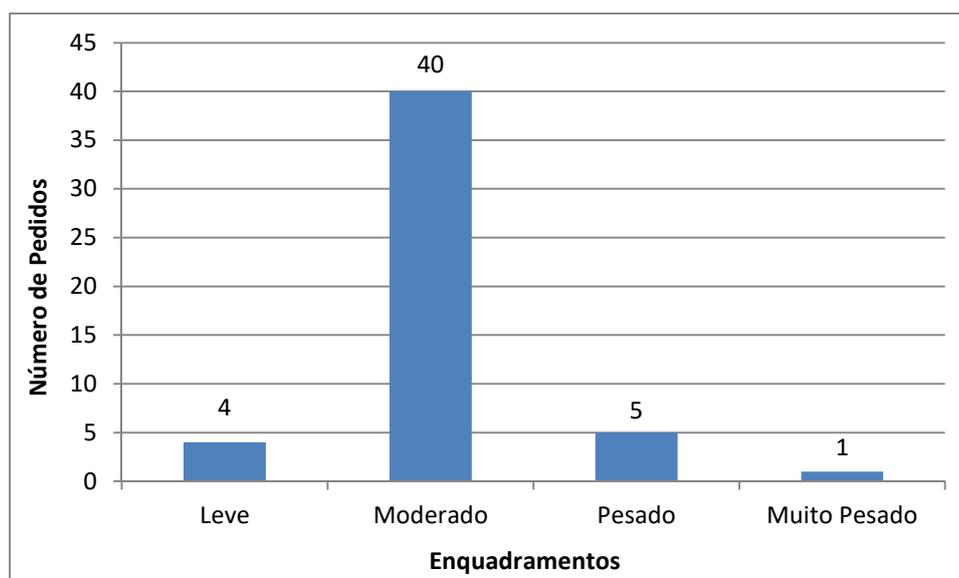


Figura 58 – Enquadramentos por IGC_{Y1} dos pedidos da amostra

Da análise das Figuras 57 e 58, verifica-se que nenhum dos pedidos de patente da amostra foi enquadrado como muito leve. A maioria dos pedidos foi enquadrada como

moderado (80%), em segundo lugar como pesado (10%), seguidos pelos leves (8%) e, por fim, os muito pesados, que foram apenas um (2%). Vale ressaltar novamente que este resultado se mostra coerente com o dos testes anteriores e, assim como no caso da amostra padrão com tempo, neste caso a amostra de dados de IGC_{Y1} também apresentou distribuição normal.

As Tabelas 82 a 84 detalham os dados dos pedidos enquadrados como muito pesados, pesados e leves com relação ao IGC_{Y1} .

Tabela 82 – Pedido de patente enquadrado como muito pesado por IGC_{Y1}

Pedido Muito Pesado no Enquadramento por IGC_{Y1}							
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
28	179	12	43	234	5	44	49

Fonte: Elaboração Própria.

Tabela 83 – Pedidos de patente enquadrados como pesados por IGC_{Y1}

Pedidos Pesados no Enquadramento por IGC_{Y1}							
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
1	125	15	0	140	6	23	29
8	62	6	6	74	7	29	36
13	63	9	9	81	14	36	50
20	61	14	28	103	4	40	44
47	239	12	3	254	7	24	31

Fonte: Elaboração Própria.

Tabela 84 – Pedidos de patente enquadrados como leves por IGC_{Y1}

Pedidos Leves no Enquadramento por IGC_{Y1}							
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
10	19	1	6	26	3	4	7
12	1	1	1	3	1	2	3
19	23	1	1	25	1	6	7
23	22	1	4	27	2	4	6

Fonte: Elaboração Própria.

Como pode ser observado das Tabelas 82 a 84, conforme esperado os enquadramentos pelo IGC_{Y1} mostram um padrão específico quanto às três variáveis do caso simulado: número de páginas de reivindicações, número total de páginas e número total de reivindicações.

De uma análise específica da Tabela 82, verifica-se que o pedido 28 apresentou número de páginas de reivindicações (12), número total de páginas (234) e número total de reivindicações (49) muito acima das respectivas médias (4,08, 72,42 e 16,18). Consequentemente, o pedido foi enquadrado como muito pesado.

Considerações análogas às anteriores podem ser feitas para os pedidos enquadrados como pesados, conforme ilustra a Tabela 83. Cabe destacar que todas as três variáveis relevantes dos pedidos 1, 8, 13, 20 e 47 apresentaram valores acima da média geral, sendo pelo menos duas delas razoavelmente acima da média. Todavia, tais valores não foram tão acima da média (mais do que três desvios padrão) a ponto de gerar enquadramentos muito pesados e, sendo assim, o modelo sugeriu um enquadramento como pesado.

Por fim, da análise da Tabela 72 também pode ser observado um padrão para as três variáveis avaliadas. Os pedidos 10, 19 e 23 apresentaram as três variáveis com valores razoavelmente abaixo da média. Todavia, os valores não foram tão abaixo da média a ponto de gerar algum enquadramento muito leve e, sendo assim, o modelo os enquadrado como leves. Especificamente com relação ao pedido 12, verifica-se que ele é bem pequeno, com apenas três páginas e três reivindicações, sendo uma delas independente. Ressalte-se que os percentis 5 para as variáveis número de páginas de reivindicações, número total de páginas e número total de reivindicações foram, respectivamente, iguais a 1, 22 e 4,9. Nota-se que o pedido está no limite do percentil 5 quanto ao número de páginas de reivindicações, mas bem abaixo do limite para o total de páginas e de reivindicações. Conseqüentemente, este pedido deveria ser enquadrado como muito leve, e a análise de percentil 5 pode corrigir enquadramento e auxiliar na decisão.

5.4.3.3 Análise das Subclasses da CIP da Amostra de Redistribuição

5.4.3.3.1 ZAE Original dos Examinadores da Amostra de Redistribuição

A Tabela 85 a seguir mostra a configuração das zonas de atuação de cada um dos 5 examinadores por subclasses da CIP para a distribuição original. Destaca-se que a área cinza corresponde às subclasses das Zonas de Atuação Específicas (ZAE) dos examinadores.

Tabela 85 – Zonas de Atuação dos Examinadores por Subclasses da CIP

ZAE original da Amostra de Redistribuição									
1		2		3		4		5	
Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%
A61K	60,00%	A61K	40,00%	A61K	60,00%	A61K	50,00%	A61K	60,00%
C07K	30,00%	G01N	30,00%	G01N	30,00%	C07K	40,00%	C07K	20,00%
C12N	10,00%	C12Q	10,00%	B82B	10,00%	C12N	10,00%	C12Q	20,00%
		C07K	10,00%						
		C12N	10,00%						

Fonte: Elaboração Própria

Da análise da Tabela 85, verifica-se que as configurações de distribuição nas ZAE de cada um dos examinadores são bem concentradas. Enquanto os examinadores 1, 3 e 4 apresentam nove dos seus dez pedidos distribuídos dentro das ZAE de referência, o examinador 2 apresenta sete e o examinador 5 apresenta oito. Tal fato evidencia uma tentativa de distribuição de carga atual utilizando como base a especialidade/áreas de conhecimento dos examinadores, sendo que a amostra avaliada parece contribuir para tal, dado que apresentou variedade de pedidos com subclasses das ZAE para distribuição.

5.4.3.3.2 Configuração Original dos Pedidos

A Tabela 86 mostra as medianas das variáveis dos pedidos com os dados separados por examinador. Nesta tabela foram hachurados: na cor cinza claro as variáveis dos pedidos de examinadores com mediana abaixo da mediana geral dos dados; e na cor cinza escuro as variáveis dos pedidos dos examinadores com mediana acima da mediana geral dos dados.

Tabela 86 – Medianas da Amostra de Redistribuição separadas por examinador

Medianas							
Examinador	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
1	68,5	3,5	5,5	77,5	4,5	11,5	15,5
2	44,5	3,5	10,5	57	3	8,5	13
3	34,5	3	5	48	2	11,5	14,5
4	36	2,5	7,5	50,5	2,5	10,5	13
5	41,5	3	5,5	55	5	8,5	13,5
Divisão	40,5	3,0	6,0	52,5	3,0	10,5	14,0

Fonte: Elaboração Própria.

Da análise da Tabela 86, verifica-se que praticamente todos os examinadores apresentam, para pelo menos duas das três variáveis em análise, medianas tanto abaixo quanto acima da mediana geral dos dados da divisão. Nota-se ainda que o examinador 1 apresenta as medianas das três variáveis acima das respectivas medianas gerais da divisão, sendo que o número total de páginas bem acima, sugerindo que o mesmo pode estar levemente sobrecarregado em relação aos demais. Por outro lado, o examinador 4 apresenta as medianas das três variáveis em análise abaixo da mediana geral da divisão, sugerindo que possa estar com a carga um pouco leve com relação aos demais. Em suma, a *priori* o balanceamento parece estar bom, com alguns desvios a tentar corrigir ou minimizar com uma nova redistribuição.

5.4.3.3.3 Redistribuição dos Pedidos

5.4.3.3.3.1 Nova Configuração das ZAE dos Examinadores

A Tabela 87 mostra a nova configuração das ZAE dos examinadores após ser realizada uma distribuição com a nova lógica desenvolvida.

Tabela 87 – ZAE dos examinadores após a nova distribuição

Novas ZAE da Amostra de Redistribuição - IGC _{v1}									
1		2		3		4		5	
Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%
A61K	60,00%	A61K	40,00%	A61K	70,00%	C07K	60,00%	A61K	60,00%
C07K	20,00%	G01N	30,00%	G01N	30,00%	A61K	40,00%	C07K	20,00%
C12N	20,00%	C12Q	10,00%					C12Q	20,00%
		C12N	10,00%						
		B82B	10,00%						

Fonte: Elaboração Própria

A Tabela 88 mostra uma comparação das porcentagens de pedidos dentro das ZAE em relação ao total de pedidos distribuídos para cada examinador nas distribuições nova e original.

Tabela 88 – Porcentagem de pedidos dentro das ZAE antes e após a nova distribuição

Examinador	1	2	3	4	5
Nova ZAE	80,00%	70,00%	100,00%	100,00%	80,00%
ZAE Original	90,00%	70,00%	90,00%	90,00%	80,00%

Fonte: Elaboração Própria

Da análise das Tabelas 87 e 88, verifica-se que o equilíbrio na concentração de exame dentro das ZAE de cada um dos examinadores que já era muito bom (vide tabela 85), foi ainda melhorado, dado que agora dois examinadores (3 e 4) estão com todos os dez pedidos dentro das suas respectivas ZAE e o examinador 2 continua sendo o único com a menor concentração de pedidos dentro da sua ZAE (70%).

5.4.3.3.3.2 Nova Configuração dos Pedidos Redistribuídos

A Tabela 89 mostra as medianas das variáveis dos pedidos redistribuídos com os dados separados por examinador. Nesta tabela foram hachurados: na cor cinza claro as variáveis dos pedidos de examinadores com mediana abaixo da mediana geral dos dados; e na

cor cinza escuro as variáveis dos pedidos dos examinadores com mediana acima da mediana geral dos dados.

Tabela 89 – Medianas da Amostra de Redistribuição separadas por examinador

medianas							
Examinador	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
1	35,5	3,5	5,5	46,5	3,5	11	13
2	42	3	8	53	2	10	14,5
3	31	3	4,5	54,5	3	11	14,5
4	32,5	2,5	4,5	48	3	9	11,5
5	46	3	6	57,5	3,5	9	14,5
Geral	40,5	3,0	6,0	52,5	3,0	10,5	14,0

Fonte: Elaboração Própria.

Da análise da Tabela 89, verifica-se que ainda há algumas variáveis de pedidos dos examinadores com mediana tanto abaixo quanto acima da mediana geral dos dados da divisão. Todavia, nota-se que agora estes desvios foram reduzidos, principalmente com relação às variáveis número de páginas de reivindicações e número total de reivindicações. Embora com diferenças pequenas para duas das três variáveis, o examinador 4 apresentou medianas das três variáveis abaixo da mediana geral, sugerindo que possa estar com carga um pouco leve em relação aos demais. Por outro lado, nenhum dos examinadores apresentou medianas das três variáveis acima da mediana geral. Estes resultados evidenciam que, além de melhorar um pouco o equilíbrio das ZAE (vide tabelas 87 e 88), a nova redistribuição teve um efeito positivo no ajuste fino das medianas das amostras dos examinadores, algo que será quantificado por meio do cálculo do IBD e avaliado por testes de hipótese.

5.4.3.3.3.3 Índice de Balanceamento da Distribuição (IBD) antes e após a Redistribuição

Foram calculados os índices IBD da distribuição original da amostra e da sua redistribuição. Para realização dos cálculos foram utilizadas as medianas constantes da tabela 86 (distribuição original) e da tabela 89 (redistribuição).

Por meio da equação (16), foram então obtidos IBD igual a 0,87 para o caso original e IBD igual a 0,91 para a redistribuição, ou seja, mesmo em uma divisão com distribuição já muito boa (IBD original muito alto), ainda houve um aumento de 4,6 % do IBD com nova distribuição, demonstrando que as medianas dos pedidos dos examinadores após a redistribuição se aproximam mais da mediana geral da divisão. Tal fato corrobora que com a nova distribuição temos uma tendência de maior equilíbrio quanto ao volume de dados e tempo/esforço dos pedidos distribuídos aos examinadores.

5.4.3.3.4 Testes de Hipótese

A Tabela 90 mostra os resultados de p-valor dos testes de Mann-Whitney realizados para cada uma das três variáveis e cada um dos cinco examinadores.

Tabela 90 – Resultados de p-valor dos Testes de Mann-Whitney

p-valor dos testes de Mann-Whitney			
Examinador	Páginas de Reivindicações	Total de Reivindicações	Total de Páginas
1	0,638	0,728	0,818
2	0,912	0,841	0,992
3	0,984	0,596	0,638
4	0,435	0,490	0,638
5	0,826	0,984	0,472

Fonte: Elaboração Própria.

Da análise da Tabela 90, verifica-se que as amostras de todas as três variáveis de todos os cinco examinadores passaram nos testes. Estes resultados evidenciam que as medianas das amostras de cada um dos examinadores podem ser consideradas iguais a mediana da amostra geral da divisão. Vale ressaltar ainda a maioria de valores elevados de p-valor, ou seja, os resultados passaram muito bem nos testes, sendo que o examinador que apresentou os menores valores foi o examinador 4, sendo coerente com os resultados e análises da tabela 89.

5.4.4 Amostra de Redistribuição – Área de Mecânica

A Tabela 91 mostra, para o conjunto total de 50 pedidos, um quadro geral das principais variáveis obtidas e os resultados dos cálculos de estatísticas básicas tais como: média, desvio padrão, mediana, moda, mínimo, máximo e coeficiente de variação.

Tabela 91 – Estatísticas Básicas da Amostra de Redistribuição

	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Reivindicações Independentes	Reivindicações Dependentes	Total de Subclasses	Ano de Depósito	Inventores	Prioridades
Média	14,88	3,30	6,90	1,54	9,68	1,54	10,16	2,56	1,02
Desvio Padrão	10,49	1,92	6,11	0,81	5,75	0,73	1,15	1,86	0,59
Mediana	11,50	3,00	5,00	1,00	9,00	1,00	10,00	2,00	1,00
Moda	7	2	6	1	9	1	10	1	1
Mínimo	4	1	0	1	0	1	6	1	0
Máximo	47	10	29	4	28	4	13	7	3
Coeficiente de Variação	70,52%	58,16%	88,48%	52,82%	59,38%	47,68%	11,31%	72,83%	57,71%

Fonte: Elaboração Própria.

Da análise da Tabela 91, verifica-se que, para quase todas as variáveis do pedido a serem consideradas, os desvios padrão (e variância, pois é o quadrado do desvio) e coeficientes de variação (relação entre o desvio padrão e a média) são altos. A exceção é o

ano de depósito que apresenta coeficiente de variação baixo (11,31%) e moda igual a 10, evidenciando que a distribuição original para os examinadores ficou concentrada em poucos anos de depósito, principalmente no ano de 2010. Adicionalmente, modas, médias e medianas da maioria das variáveis são distintas entre si, sendo que os valores mínimos e máximos são bastante diferentes. Tais resultados sugerem que alta amplitude e boa dispersão são características intrínsecas dos pedidos analisados, ou seja, há uma variabilidade razoável entre eles mesmo considerando pedidos de mesmo tipo (6.21 com aproveitamento de buscas), sendo que, conforme desejado, o perfil dos dados é parecido com o da amostra padrão com tempo e com o das amostras de redistribuição das demais áreas já analisadas.

5.4.4.1 Execução da Análise de Componentes Principais: Caso 3 Var (2)

A Tabela 92 mostra os principais resultados da execução da ACP para os dados da área de mecânica.

Tabela 92 – Autovalores e Contribuições de Variância das Componentes Principais

Componente Principal	Autovalores	Variância	Variância Acumulada
Y1	2,15	72%	72%
Y2	0,60	20%	92%
Y3	0,25	8%	100%

Fonte: Elaboração Própria.

Da análise da Tabela 92, verifica-se que no caso de utilização do critério considerando autovalores maiores do que um, ou mesmo o critério de variância acumulada (acima de 60%), selecionaríamos somente a primeira componente para a determinação do IGC, ou seja, $IGC_{\lambda > 1} = IGC_{Y1}$. As Figuras 59 e 60 mostram, respectivamente, os autovalores e as variâncias acumuladas para cada um dos casos 3 var (2), com amostras de 10 até 50 pedidos.

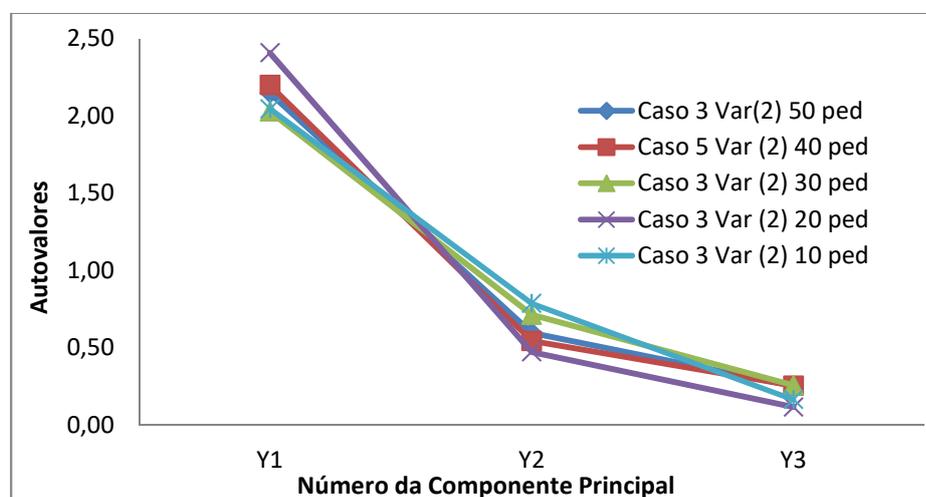


Figura 59 – Caso 3 Var (2): autovalores para amostras de 10 até 50 pedidos

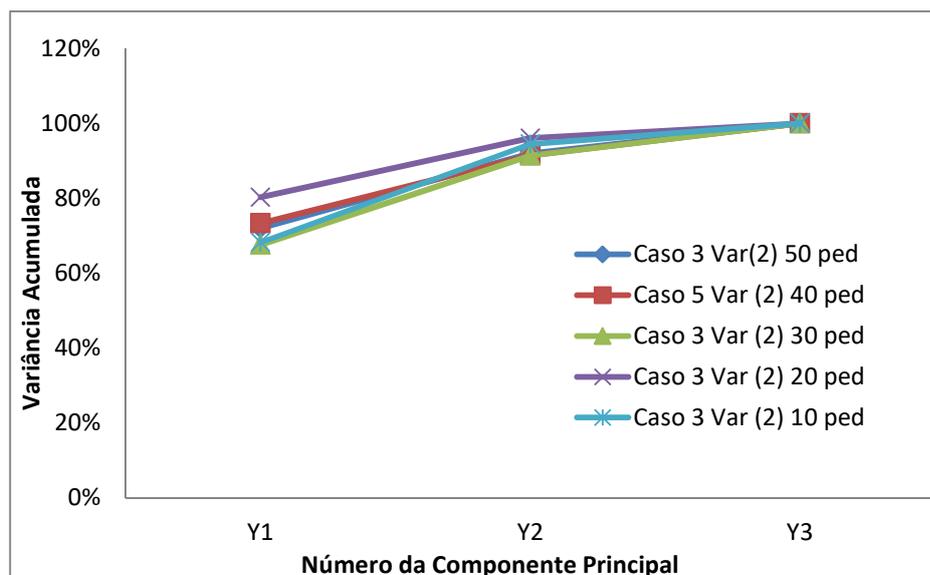


Figura 60 – Caso 3 Var (2): variância acumulada para amostras de 10 até 50 pedidos

Da análise da Figura 59, verifica-se que somente o caso com amostra de 20 pedidos se desvia um pouco dos demais no autovalor da componente Y1 e, ainda assim, com um resultado melhor que os demais (autovalor maior). O mesmo fenômeno pode ser observado ao analisar as variâncias acumuladas das amostras conforme ilustrado na Figura 60. Estes resultados reforçam que, para os casos com três variáveis que se pretenda executar, já é possível utilizar uma amostra de 10 pedidos de patente, sendo preferível a partir de 20 pedidos, de modo que se consiga um melhor rendimento do método de ACP.

A Tabela 93 mostra os escores das duas primeiras componentes principais para cada uma das variáveis consideradas do caso 3 var (2) com 50 pedidos.

Tabela 93 – Escores das Componentes Principais de cada variável

Variáveis	Coeficiente de Ponderação	
	Y1	Y2
Total de Páginas	0,5301	-0,7896
Total de Reivindicações	0,5724	0,6022
Páginas de Reivindicações	0,6256	0,1174

Fonte: Elaboração Própria.

Na Tabela 93 foram hachurados na cor cinza os valores considerados significativos dos fatores para cada uma das variáveis (próximos ou acima de 0,4). Em uma primeira análise, nota-se que a componente Y1 é bastante destacada das demais, sendo responsável por explicar 72% de toda a variância. Neste caso, ambas as componentes Y1 e Y2 estão associadas às principais variáveis diretamente relacionadas ao volume de dados (total de

páginas, páginas de reivindicações, além de número de reivindicações). Este resultado é bem semelhante aos demais casos 3 var (2) já executados.

5.4.4.2 Enquadramento dos Pedidos de Patente da Amostra de Redistribuição por IGC_{Y1}

A Figura 61 mostra os IGC_{Y1} padronizados (eixo vertical) para todos os cinquenta pedidos (eixo horizontal) da amostra. As retas horizontais facilitam a verificação visual dos enquadramentos. Já a Figura 62 mostra a quantidade de enquadramentos por IGC_{Y1} nas classes.

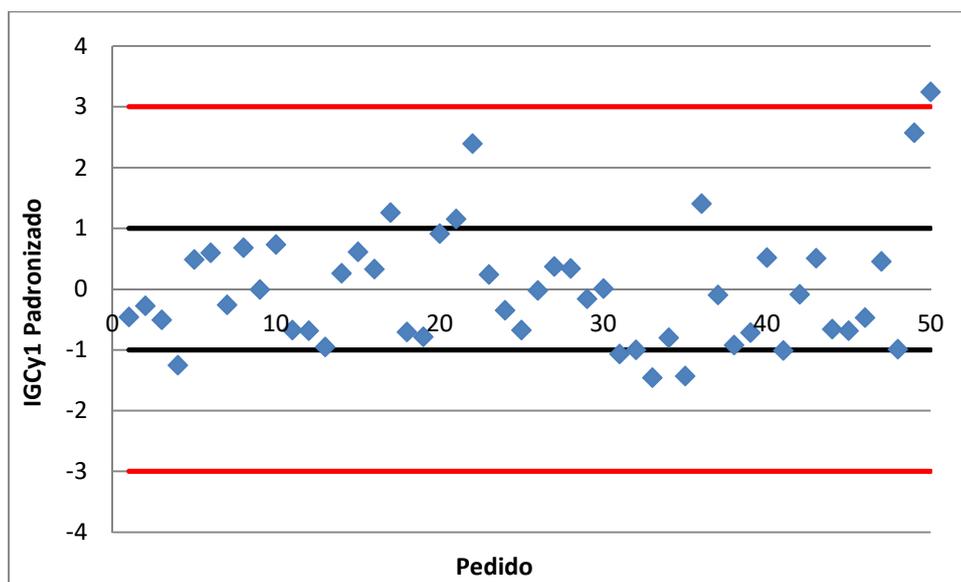


Figura 61 – IGC_{Y1} padronizado dos pedidos

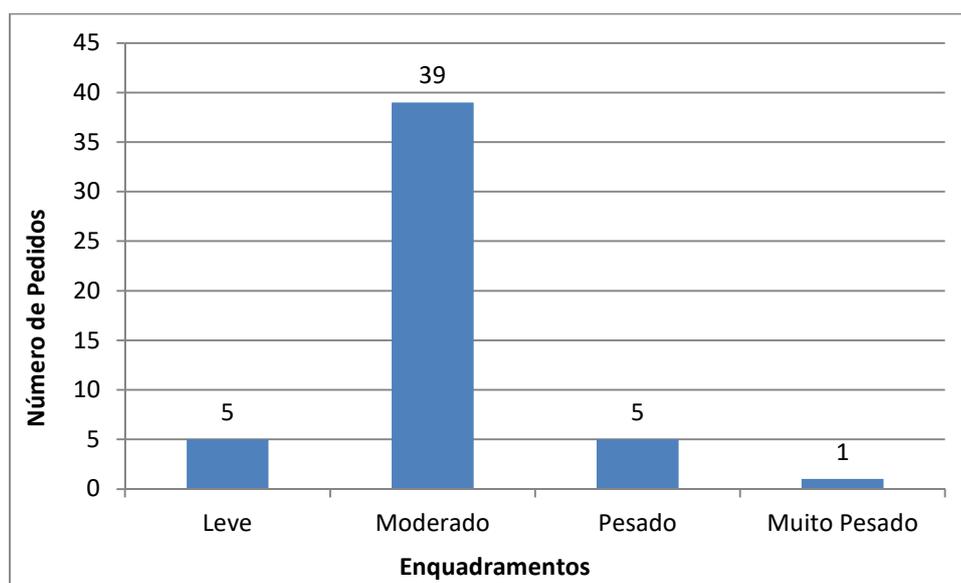


Figura 62 – Enquadramentos por IGC_{Y1} dos pedidos da amostra

Da análise das Figuras 61 e 62, verifica-se que nenhum dos pedidos de patente da amostra foi enquadrado como muito leve. A maioria dos pedidos foi enquadrada como moderado (78%), em segundo lugar pesados e leves empatados (10%) e, por fim, os muito pesados, que foram apenas um (2%). Vale ressaltar novamente que este resultado se mostra coerente com o dos testes anteriores e, assim como no caso da amostra padrão com tempo, neste caso a amostra de dados de IGC_{Y1} também apresentou distribuição normal.

As Tabelas 94 a 96 detalham os dados dos pedidos enquadrados como muito pesados, pesados e leves com relação ao IGC_{Y1} .

Tabela 94 – Pedido de patente enquadrado como muito pesado por IGC_{Y1}

Pedido Muito Pesado no Enquadramento por IGC_{Y1}							
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
50	39	8	29	76	3	25	28

Fonte: Elaboração Própria.

Tabela 95 – Pedidos de patente enquadrados como pesados por IGC_{Y1}

Pedidos Pesados no Enquadramento por IGC_{Y1}							
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
17	39	4	26	69	1	12	13
21	27	6	9	42	2	12	14
22	47	10	6	63	3	9	12
36	19	7	6	32	2	16	18
49	20	8	7	35	4	28	32

Fonte: Elaboração Própria.

Tabela 96 – Pedidos de patente enquadrados como leves por IGC_{Y1}

Pedidos Leves no Enquadramento por IGC_{Y1}							
Pedido	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
4	8	2	1	11	1	0	1
31	4	2	2	8	1	4	5
33	5	1	3	9	1	1	2
35	6	1	3	10	1	1	2
41	6	1	5	12	1	7	8

Fonte: Elaboração Própria.

Como pode ser observado das Tabelas 94 a 96, conforme esperado os enquadramentos pelo IGC_{Y1} mostram um padrão específico quanto às três variáveis do caso simulado: número de páginas de reivindicações, número total de páginas e número total de reivindicações.

De uma análise específica da Tabela 94, verifica-se que o pedido 50 apresentou número de páginas de reivindicações (8), número total de páginas (76) e número total de reivindicações (28) muito acima das respectivas médias (3,3, 25,08 e 11,22). Consequentemente, o pedido foi enquadrado como muito pesado.

Considerações análogas às anteriores podem ser feitas para os pedidos enquadrados como pesados, conforme ilustra a Tabela 95. Cabe destacar que todas as três variáveis relevantes dos pedidos 17, 21, 22, 36 e 49 apresentaram valores acima da média geral, sendo pelo menos duas delas razoavelmente acima da média. Todavia, para pelo menos uma ou duas das variáveis, tais valores não foram tão acima da média e, sendo assim, o modelo sugeriu um enquadramento como pesado.

Por fim, da análise da Tabela 96 também pode ser observado um padrão para as três variáveis avaliadas. Os pedidos 4, 31, 33, 35 e 41 apresentaram as três variáveis com valores razoavelmente abaixo da média. Todavia, os valores não foram tão abaixo da média a ponto de gerar algum enquadramento muito leve e, sendo assim, o modelo os enquadrado como leves. Vale destacar que o pedido 4, embora apresente somente uma reivindicação, tem duas páginas de quadro reivindicatório, ou seja, uma única reivindicação muito grande, sendo um pedido de difícil avaliação *a priori*, sem uma análise do seu conteúdo. Todavia, poderia ser enquadrado como muito leve. Especificamente com relação aos pedidos 33 e 35, verifica-se que eles são bem pequenos, com apenas uma página de quadro reivindicatório e duas reivindicações, sendo uma delas independente. Neste caso, os pedidos também poderiam ser enquadrados como muito leves ao invés de leves, evidenciando que, para situações em que as três variáveis apresentem coeficiente de variação muito alto, o modelo pode apresentar limitações para enquadrar pedidos muito leves. Sendo assim, mais uma vez cabe destacar que é importante haver possibilidade da realização de um ajuste fino pelo gestor responsável ou, em caso de utilização do modelo para distribuição automática, haver possibilidade de interação entre o examinador e o gestor visando a possíveis alterações consensuais.

5.4.4.3 Análise das Subclasses da CIP da Amostra de Redistribuição

5.4.4.3.1 ZAE Original dos Examinadores da Amostra de Redistribuição

A Tabela 97 a seguir mostra a configuração das zonas de atuação de cada um dos 5 examinadores por subclasses da CIP para a distribuição original. Destaca-se que a área cinza corresponde às subclasses das Zonas de Atuação Específicas (ZAE) dos examinadores.

Tabela 97 – Zonas de Atuação dos Examinadores por Subclasses da CIP

ZAE original da Amostra de Redistribuição									
1		2		3		4		5	
Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%
B63B	100,00%	F03D	70,00%	F23D	30,00%	B60T	30,00%	F04B	10,00%
		F03B	20,00%	B60K	30,00%	F16D	20,00%	F04C	10,00%
		B61B	10,00%	F23C	30,00%	B61D	10,00%	F01K	10,00%
				F02M	10,00%	B61K	10,00%	B60K	10,00%
						B60G	10,00%	F04D	10,00%
						B61J	10,00%	F21L	10,00%
						F16H	10,00%	B60D	10,00%
								B66F	10,00%
								B63B	10,00%
								F25B	10,00%

Fonte: Elaboração Própria

Da análise da Tabela 97, verifica-se que as configurações de distribuição nas ZAE dos examinadores 1 a 4 são razoavelmente concentradas. Enquanto os examinadores 2 e 3 apresentam nove dos seus dez pedidos distribuídos dentro das ZAE de referência, o examinador 1 apresenta todos os dez e o examinador 4 apresenta cinco. Tal fato evidencia uma tentativa de distribuição de carga atual utilizando como base a especialidade/áreas de conhecimento dos examinadores. Cabe um destaque para o examinador 5, dado que todos os pedidos distribuídos a ele foram de subclasses distintas, ou seja, não foi possível identificar uma ZAE específica. Tal fato pode decorrer do dito examinador ser mais versátil que os demais.

5.4.4.3.2 Configuração Original dos Pedidos

A Tabela 98 mostra as medianas das variáveis dos pedidos com os dados separados por examinador. Nesta tabela foram hachurados: na cor cinza claro as variáveis dos pedidos de examinadores com mediana abaixo da mediana geral dos dados; e na cor cinza escuro as variáveis dos pedidos dos examinadores com mediana acima da mediana geral dos dados.

Tabela 98 – Medianas da Amostra de Redistribuição separadas por examinador

Medianas							
Examinador	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
1	16,5	3,5	5,5	25,5	2	10,5	12
2	8,5	2	3	16,5	1	4	5,5
3	8,5	2	7	19	1	9	10
4	14	2	5,5	25,5	1	10	11
5	13	3	4,5	21,5	2	10,5	12
Divisão	11,5	3,0	5,0	20,5	1,0	9,0	10,5

Fonte: Elaboração Própria.

Da análise da Tabela 98, verifica-se que praticamente todos os examinadores apresentam, para as três variáveis em análise, medianas tanto abaixo quanto acima da mediana geral dos dados da divisão. Nota-se ainda que o examinador 1 apresenta as medianas das três variáveis acima das respectivas medianas gerais da divisão, sendo que o número total de páginas bem acima, sugerindo que o mesmo pode estar levemente sobrecarregado em relação aos demais. Por outro lado, os examinadores 2 e 3 apresentam as medianas das três variáveis em análise abaixo da mediana geral da divisão, sugerindo que possam estar com a carga um pouco leve com relação aos demais. Destaque para o examinador 2 em que tais valores são razoavelmente abaixo das medianas gerais da divisão. Em suma, *a priori* o balanceamento parece estar de regular a bom, com alguns desvios a tentar corrigir ou minimizar com uma nova redistribuição.

5.4.4.3.3 Redistribuição dos Pedidos

5.4.4.3.3.1 Nova Configuração das ZAE dos Examinadores

A Tabela 99 mostra a nova configuração das ZAE dos examinadores após ser realizada uma distribuição com a nova lógica desenvolvida.

Tabela 99 – ZAE dos examinadores após a nova distribuição

Novas ZAE da Amostra de Redistribuição - IGC _{v1}									
1		2		3		4		5	
Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%	Subclasse	%
B63B	90,00%	F03D	70,00%	F23D	30,00%	B60T	30,00%	B63B	20,00%
F25B	10,00%	F03B	20,00%	F23C	30,00%	F16D	20,00%	F02M	10,00%
		B61K	10,00%	B60K	20,00%	F04B	10,00%	B60G	10,00%
				B61D	10,00%	B61B	10,00%	F16H	10,00%
				B61J	10,00%	B60K	10,00%	F04C	10,00%
						F04D	10,00%	F01K	10,00%
						B60D	10,00%	B60K	10,00%
								F21L	10,00%
								B66F	10,00%

Fonte: Elaboração Própria

A Tabela 100 mostra uma comparação das porcentagens de pedidos dentro das ZAE em relação ao total de pedidos distribuídos para cada examinador nas distribuições nova e original.

Tabela 100 – Porcentagem de pedidos dentro das ZAE antes e após a nova distribuição

Examinador	1	2	3	4	5
Nova ZAE	90,00%	90,00%	80,00%	50,00%	0,00%
ZAE Original	100,00%	90,00%	90,00%	50,00%	0,00%

Fonte: Elaboração Própria

Da análise das Tabelas 99 e 100, verifica-se que o equilíbrio na concentração de exame dentro das ZAE de cada um dos examinadores que já era bom (vide tabela 97), foi praticamente mantido. Os examinadores 1 e 3 que antes apresentavam, respectivamente, dez e nove pedidos dentro de suas ZAE, agora apresentam nove e oito. Cabe destacar que tais alterações não comprometem as concentrações das ZAE, que permanecem boas, sendo que, como será mostrado adiante, tais alterações foram necessárias para obter um maior equilíbrio da distribuição quanto ao volume de dados/complexidade dos pedidos.

5.4.4.3.2 Nova Configuração dos Pedidos Redistribuídos

A Tabela 101 mostra as medianas das variáveis dos pedidos redistribuídos com os dados separados por examinador. Nesta tabela foram hachurados: na cor cinza claro as variáveis dos pedidos de examinadores com mediana abaixo da mediana geral dos dados; e na cor cinza escuro as variáveis dos pedidos dos examinadores com mediana acima da mediana geral dos dados.

Tabela 101 – Medianas da Amostra de Redistribuição separadas por examinador

Examinador	medianas						
	Páginas de Relatório	Páginas de Reivindicações	Páginas de Desenhos	Total de Páginas	Reivindicações Independentes	Reivindicações Dependentes	Total de Reivindicações
1	15	3,5	6	22,5	2	10,5	12,5
2	8,5	2	6,5	19	1	9	10
3	11	3,5	4,5	16,5	2	9,5	12
4	11	2	4,5	17	1	9	10,5
5	12,5	2,5	5,5	21	1	9,5	10,5
Geral	11,5	3,0	5,0	20,5	1,0	9,0	10,5

Fonte: Elaboração Própria.

Da análise da Tabela 101, verifica-se que ainda há algumas variáveis de pedidos dos examinadores com mediana tanto abaixo quanto acima da mediana geral dos dados da divisão. Todavia, nota-se que agora estes desvios foram reduzidos em todas as três variáveis em análise. Embora com diferenças pequenas para duas das três variáveis, o examinador 1 apresentou medianas das três variáveis acima da mediana geral, sugerindo que possa estar levemente sobrecarregado em relação aos demais. Por outro lado, o examinador 2 apresentou medianas das três variáveis abaixo da mediana geral, porém duas delas com diferenças bem pequenas, não sendo possível definir que esteja com carga leve. Estes resultados evidenciam que, além de manter o bom equilíbrio das ZAE (vide tabelas 99 e 100), a nova redistribuição teve um efeito positivo no ajuste fino das medianas das amostras dos examinadores (vide

tabelas 98 e 101), algo que será quantificado por meio do cálculo do IBD e avaliado por testes de hipótese.

5.4.4.3.3 Índice de Balanceamento da Distribuição (IBD) antes e após a Redistribuição

Foram calculados os índices IBD da distribuição original da amostra e da sua redistribuição. Para realização dos cálculos foram utilizadas as medianas constantes da tabela 98 (distribuição original) e da tabela 101 (redistribuição).

Por meio da equação (16), foram então obtidos IBD igual a 0,81 para o caso original e IBD igual a 0,85 para a redistribuição, ou seja, um aumento de 5% do IBD com nova distribuição, demonstrando que as medianas dos pedidos dos examinadores após a redistribuição se aproximam mais da mediana geral da divisão. Tal fato corrobora que com a nova distribuição temos uma tendência de maior equilíbrio quanto ao volume de dados e tempo/esforço dos pedidos distribuídos aos examinadores.

5.4.4.3.4 Testes de Hipótese

A Tabela 102 mostra os resultados de p-valor dos testes de Mann-Whitney realizados para cada uma das três variáveis e cada um dos cinco examinadores.

Tabela 102 – Resultados de p-valor dos Testes de Mann-Whitney

p-valor dos testes de Mann-Whitney			
Examinador	Páginas de Reivindicações	Total de Reivindicações	Total de Páginas
1	0,258	0,347	0,379
2	0,589	0,726	0,749
3	0,447	0,944	0,610
4	0,453	0,928	0,542
5	0,549	0,682	0,569

Fonte: Elaboração Própria.

Da análise da Tabela 102, verifica-se que as amostras de todas as três variáveis de todos os cinco examinadores passaram nos testes. Estes resultados evidenciam que as medianas das amostras de cada um dos examinadores podem ser consideradas iguais a mediana da amostra geral da divisão. Vale ressaltar ainda a maioria de valores elevados de p-valor, ou seja, os resultados passaram muito bem nos testes, sendo que o examinador que apresentou os menores valores foi o examinador 1, o que é coerente com os resultados e análises da tabela 101.

5.4.5 Comparação de Áreas Técnicas - Amostras de Redistribuição

Primeiramente destaque-se que, para as quatro áreas testadas, de maneira idêntica aos dados de eletricidade, as amostras contendo 20 ou mais pedidos passaram nos testes de Bartlett, sendo que apenas amostras com 10 pedidos não foram aprovadas. Além disso, das análises gráficas das variâncias acumuladas das componentes principais, verificou-se resultados similares, ou seja, estabilidade de comportamento das componentes principais da ACP para todos os tamanhos de amostras testados, principalmente a partir de amostras com 20 ou mais pedidos. A Figura 63 mostra os gráficos de comportamento das variâncias acumuladas para todas as áreas avaliadas.

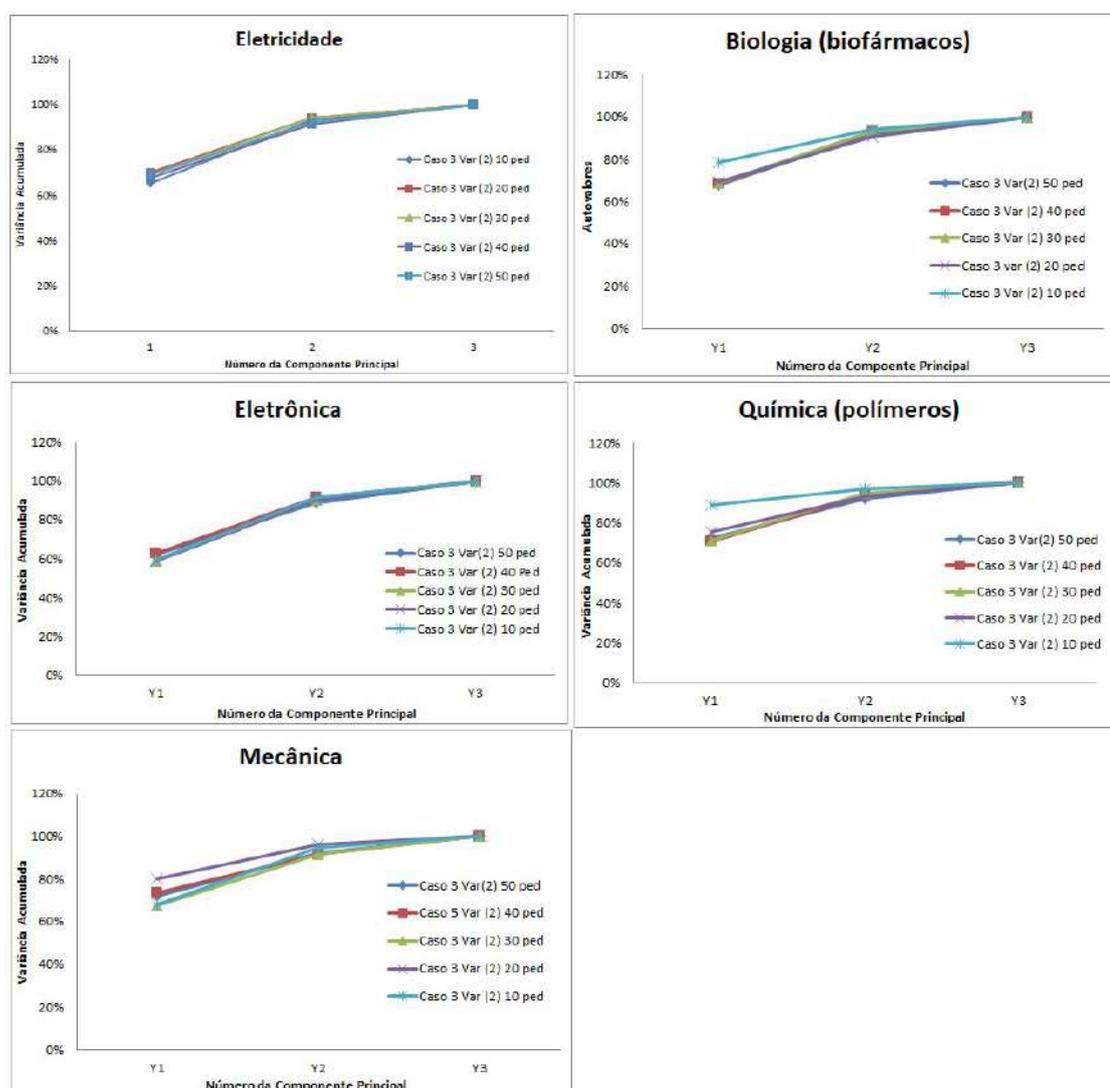


Figura 63 – Comportamento das variâncias acumuladas das áreas analisadas

Sendo assim, um primeiro critério sugerido para aplicação do modelo desenvolvido a dados de outras áreas é o de que, para simulação de casos com 3 variáveis, a amostra deve conter no mínimo 20 pedidos de patente. Tal critério está em consonância com a prática de pesquisadores da área, cuja recomendação é utilizar um número de amostras de no mínimo cinco vezes o número de variáveis do problema.

Adicionalmente, tendo em vista os resultados obtidos quando da validação do modelo para a área de eletricidade e os critérios de escolha de componentes principais comumente utilizados na prática, visando a uma garantia de rendimento adequado do método da ACP, além de uma boa representatividade da variância dos dados por meio do índice IGC_{Y1} , faz-se necessária uma elevada variância da componente Y1, de aproximadamente 60% ou mais. A Figura 64 a seguir mostra as variâncias da componente principal Y1 das áreas estudadas para amostras de 20 até 50 pedidos.

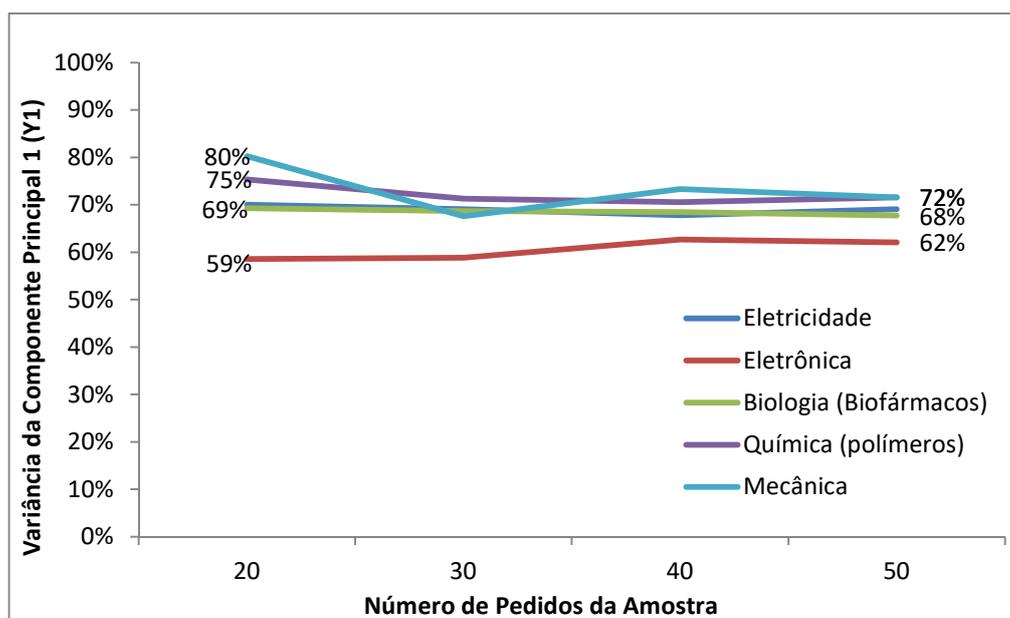


Figura 64 – Variâncias da componente principal Y1 das áreas analisadas

Da análise da figura 64 verifica-se que, ainda que com variação no tamanho das amostras, todas as áreas apresentam um comportamento bastante estável, quase que constante, da variância da componente principal Y1. Destaque-se ainda que todas as variâncias das componentes principais Y1 foram elevadas, com valores acima de 59%. Conseqüentemente, foi atendido o segundo critério sugerido para aplicação do modelo desenvolvido a dados de outras áreas de que, para simulação de casos com 3 variáveis, a componente principal Y1 deve apresentar variância mínima de aproximadamente 60%.

Além de componentes principais com elevadas variâncias, quando da validação do modelo para a área de eletricidade verificou-se que, tanto as amostras de dados do tempo dispendido de exame, quanto as amostras de dados de IGC_{Y1} calculados, para todos os casos e tamanhos de amostras testados, apresentaram como característica uma distribuição estatística normal. Cabe destacar que tal característica contribui diretamente para o bom funcionamento dos enquadramentos por classes utilizados. Conseqüentemente, uma condição necessária para o bom funcionamento do modelo e sugerida como um terceiro critério a ser atendido, é a obtenção de dados de IGC_{Y1} com distribuição estatística normal, algo que também foi verificado para todas as amostras das áreas analisadas.

Cabe destacar que, embora possa haver diferenças de procedimento de exame entre áreas, o quadro reivindicatório é parte do pedido de patente de grande ou maior relevância a ser considerada. Tal relevância aumenta ainda mais quando da realização de um exame simplificado (com aproveitamento de buscas internacionais), e pode ser analisada tanto em termos de tamanho das reivindicações quanto número de reivindicações. Neste contexto, verifica-se que o número de páginas do quadro reivindicatório é uma variável bem representativa para análise. A Figura 65 a seguir mostra as correlações do IGC_{Y1} com a variável número de páginas do quadro reivindicatório dos casos simulados para as áreas em estudo.

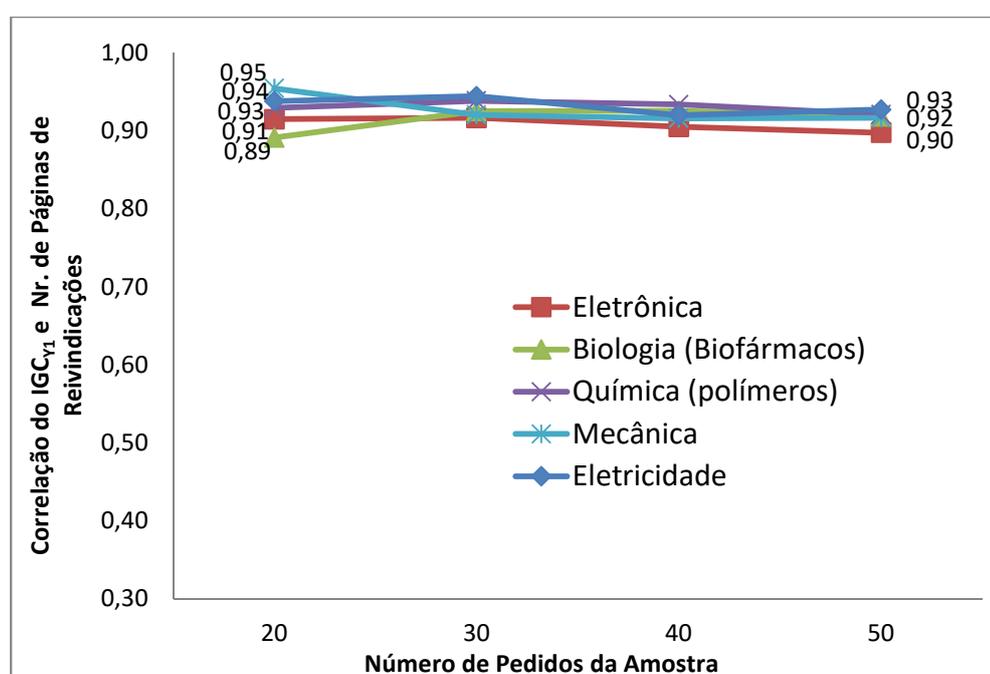


Figura 65 – Correlações do IGC_{Y1} com o número de páginas de reivindicações

Da análise da figura 65 verifica-se que, mesmo variando o tamanho das amostras, todas as áreas apresentaram um comportamento bastante estável da correlação do IGC_{Y1} com as páginas do quadro reivindicatório, quase que uma correlação constante. Destaque-se ainda que todos os valores de correlação foram muito elevados, oscilando em torno de 0,9. Estes resultados mostram que, tanto para a área de eletricidade quanto para as demais áreas, há um comportamento padrão do modelo, com clara ênfase nas páginas de reivindicações. Vale ressaltar que, conforme visto ao longo de todos os testes de validação do modelo, a variável páginas de reivindicações foi disparadamente a mais confiável, apresentando correlações com o tempo bastante estáveis e elevadas, sempre perto de 0,8. Tal comportamento é considerado importante, pois evidencia uma atuação do modelo de forma conservadora e coerente com o processo de exame, afinal, uma avaliação aleatória tende a ser menos equilibrada e justa do que uma avaliação considerando esta variável. Assim sendo, um quarto critério sugerido para aplicação do modelo desenvolvido a outras áreas é que a correlação do IGC_{Y1} com as páginas de reivindicações seja bastante elevada, com valores próximos a 0,9.

Por fim, ao aplicar o modelo tanto para a área de eletricidade quanto para as demais áreas, verificou-se que os valores de IBD foram altos (sempre acima de 0,8) e que, nos testes de redistribuição, os valores de IBD sempre aumentaram, com valores de 5 a 10% maiores, evidenciando matematicamente a eficiência da nova metodologia de distribuição. Sendo assim, é desejável que uma boa distribuição apresente IBD acima de 0,8 seguindo o padrão testado para todos os casos simulados. Obviamente, em caso de simulações de redistribuição, ou seja, comparação com distribuição passada, é esperado que o IBD aumente de valor.

Em suma, por meio das análises realizadas, e com a determinação dos critérios/condições decorrentes delas, foi possível garantir uma boa execução do método da ACP, um comportamento estatístico do IGC semelhante ao caso já validado com o tempo (com os dados da área de eletricidade) e o bom funcionamento dos enquadramentos em classes (funcionam bem para distribuição normal). Além disso, uma correlação similar com a principal variável de interesse (número de páginas de reivindicações) faz com que o modelo seja conservador, sendo que em conjunto com uma melhora do índice IBD, há uma confirmação de maior equilíbrio das redistribuições testadas. Portanto, considera-se que o modelo desenvolvido pode ser aplicado a outras áreas do INPI desde que seja realizada uma análise estatística prévia da amostra a ser utilizada e sejam atendidos os seguintes critérios mínimos: a) número mínimo de 20 amostras para simulação de casos com 3 variáveis; b) variância acumulada da componente principal $Y1$ elevada (cerca de 60% ou maior); c)

amostra de dados do IGC_{Y1} com distribuição estatística aproximadamente normal; d) alta correlação dos dados de IGC_{Y1} com o número de páginas de reivindicações (cerca de 0,90); e) valor alto de IBD nas simulações (cerca de 0,80 ou maior) sendo que, em caso de testes de redistribuição, verificar um aumento do IBD nas simulações.

6 CONCLUSÃO

Inicialmente, com a realização do presente trabalho, foram identificadas dez possíveis variáveis relevantes para avaliação e distribuição dos pedidos de patente aos examinadores. Entre essas variáveis, foram identificadas aquelas que se relacionam diretamente à voluminosidade de um documento de patente, ou seja, ao volume de dados com que o examinador tem de lidar quando do exame de patentes. Em consonância com a nossa hipótese inicial 1, são elas: número de páginas de relatório descritivo, número de páginas de reivindicações, número de páginas de desenhos, número de reivindicações independentes e número de reivindicações dependentes.

Com a aplicação do modelo de análise de componentes principais (ACP) em uma primeira amostra de dados, denominada Amostra de Teste Inicial, foi verificado que as componentes estavam coerentes com as hipóteses iniciais e, aplicados os testes que mostravam que o problema a ser resolvido apresenta dados que justificam a sua utilização (testes KMO e Bartlett), foi então proposto um índice geral de complexidade (IGC).

De posse desta amostra inicial, contendo um número grande de pedidos examinados ao longo de dois anos, foram ainda determinadas as Zonas de Atuação Específicas (ZAE) dos examinadores, aquelas subclasses da IPC (áreas tecnológicas) que eles mais examinam de acordo com sua formação acadêmica e experiência profissional. Estas ZAE são de grande relevância, pois estas subclasses são um dos critérios utilizados para distribuição dos pedidos de patente aos examinadores, sendo importante sua comparação antes e depois de qualquer redistribuição que se faça.

Os pedidos de patente foram ainda enquadrados em até cinco classes: muito leve, leve, moderado, pesado e muito pesado, sendo que os enquadramentos tiveram como referência os valores de IGC, considerando faixas equivalentes à média do índice mais um, três ou mais desvios padrão. Foi então realizada uma redistribuição dos pedidos com foco nas ZAE dos examinadores e nos enquadramentos. Os resultados mostram que as medianas dos pedidos dos examinadores ficaram mais próximas das medianas gerais da divisão, sugerindo que a nova distribuição seja mais justa quanto ao volume de dados que a original. Além disso, com a nova distribuição os examinadores tiveram uma maior parte de seus pedidos alocados dentro das suas respectivas ZAE, ou seja, examinariam mais pedidos de suas áreas específicas de

formação e preferência, sugerindo ainda que a nova distribuição contribui para uma maior eficiência, qualidade e motivação.

Adicionalmente, os resultados obtidos sugerem que, embora as cinco variáveis diretamente relacionadas ao volume de dados tendam a ser as que mais impactam no processo de exame, todas as dez variáveis selecionadas, em alguma medida, influenciam na análise dos pedidos de patente quando de uma avaliação de complexidade, fato que confirma a nossa hipótese inicial 2.

Por outro lado, como complexidade é algo relativo, para investigarmos se essa complexidade de fato capta o tempo/esforço de exame, foi realizada uma análise de sensibilidade do modelo desenvolvido de modo a verificar as correlações das variáveis e IGC com o tempo. Para isso, foi então necessário obter uma nova amostra de pedidos de patente, denominada Amostra Padrão, desta vez com a coleta adicional da variável tempo dispendido no exame.

Neste contexto, foram realizadas simulações considerando diferentes variáveis e tamanhos da amostra padrão, tendo sido aplicado o método de ACP e o modelo desenvolvido, incluindo cálculo do IGC com diferentes critérios e suas correlações com o tempo. Os resultados obtidos sugerem que, para o nosso problema específico, o IGC com maior eficiência e estabilidade foi o chamado IGC_{Y1} , ou seja, utilizando somente a primeira componente principal, aquela que é mais representativa em variância total dos dados.

Vale ressaltar ainda que o caso incluindo somente três variáveis (número de páginas de reivindicações, número total de páginas e número total de reivindicações) é o recomendado para realização de testes práticos ainda mais específicos, seja por captar as influências das principais variáveis de volume direto de dados, seja pela maior simplicidade para obtenção e coleta de dados (pois não precisa separar reivindicações independentes e dependentes), ou ainda pelas consistentes altas correlações de IGC_{Y1} com o tempo, sempre perto de 0,85.

De posse desta nova amostra com coleta do tempo de exame, os pedidos de patente foram novamente enquadrados nas cinco classes definidas: muito leve, leve, moderado, pesado e muito pesado. Tais enquadramentos foram realizados duas vezes, uma primeira vez utilizando como referência a variável tempo dispendido no exame, ou seja, o enquadramento padrão de referência, e uma segunda vez utilizando o índice IGC_{Y1} , ou seja, o enquadramento sugerido pelo modelo. Ao realizar a comparação destes enquadramentos pelo tempo com os enquadramentos do modelo (variando inclusive o tamanho das amostras e o número de variáveis), os resultados mostraram uma alta similaridade, onde o modelo enquadrrou

corretamente, na média de todos os casos simulados, 81% dos pedidos de patente analisados e, para o caso com três variáveis (caso 3Var (2)), uma média de 84,6%, apresentando sempre valores acima de 80%.

De posse do modelo matemático e de critérios de enquadramento testados com as correlações com o tempo, o próximo passo foi a realização de um primeiro teste prático de redistribuição completo. Para tanto, foi necessário coletar uma terceira e última amostra, denominada de Amostra Final de Redistribuição, amostra esta com 100 pedidos de patente, sendo 10 pedidos de 10 examinadores diferentes, todos com aproveitamento de buscas em escritórios internacionais, de modo que o perfil desta nova amostra fosse similar ao da amostra padrão, nossa referência já testada anteriormente. De posse desta nova amostra foram determinadas as principais estatísticas de tendência central das amostras por examinador e calculados os Índices de Balanceamento da Distribuição (IBD) tanto para a distribuição original, quanto para a amostra redistribuída pelo IGC_{Y1} .

Os resultados obtidos com a nova redistribuição mostraram que houve um maior equilíbrio na concentração de exame dentro das ZAE de cada um dos examinadores, sendo que nas amostras de seis dos dez examinadores analisados houve um aumento do número de pedidos distribuídos a eles e pertencentes às suas próprias ZAE. Sendo assim, os resultados sugerem que esta nova configuração contribui para que os examinadores trabalhem dentro das suas áreas de formação e conhecimento específicas e, conseqüentemente, para a sua eficiência e motivação.

Destaque-se ainda que, se na distribuição original havia três examinadores com um leve desbalanceamento de carga em relação aos demais, após a redistribuição por IGC_{Y1} , não foi identificado nenhum examinador visivelmente com carga leve ou sobrecarregado em relação aos demais. Estes resultados mostram que, além de melhorar as suas ZAE, a nova redistribuição teve um efeito positivo no ajuste fino das medianas das amostras dos examinadores, algo que foi matematicamente quantificado por meio do cálculo do IBD que, na distribuição original tinha um valor de 0,83 e, após a redistribuição, aumentou para 0,90.

Para uma última análise complementar dos resultados da nova distribuição, foram executados testes de hipótese de Mann-Whitney de modo a comparar as distribuições estatísticas e medianas das amostras de pedidos distribuídos a cada examinador com a distribuição estatística e mediana da amostra geral da divisão. Verificou-se que as amostras de todos os dez examinadores passaram nos testes, evidenciando que as medianas e as

distribuições estatísticas das amostras de cada um dos examinadores podem ser consideradas iguais à mediana e à distribuição estatística da amostra geral da divisão.

O mesmo procedimento metodológico aplicado para os dados da área de eletricidade foi também executado para outras quatro áreas técnicas distintas do INPI: eletrônica, química (polímeros), biologia (biofármacos) e mecânica. Os resultados mostraram que, mesmo para áreas técnicas bastante diferentes, e ainda que com valores absolutos distintos por variáveis, as amostras apresentaram perfis de variabilidade e correlação semelhantes aos das amostras da área de eletricidade (amostras padrão com tempo e de redistribuição), as quais foram utilizadas para a validação do modelo. Verificou-se também que, quando aplicado às amostras das demais áreas, o método da ACP funcionou com boa eficiência e de maneira análoga aos casos simulados para a área de eletricidade. Destaque-se que todas as amostras de dados de IGC_{Y1} (para todas as áreas e tamanhos de amostra testados) apresentaram distribuição estatística normal, além de correlações bem altas com a variável número de páginas de reivindicações, aqui considerada a mais relevante para o exame, principalmente em pedidos com aproveitamento de buscas, onde o exame tende a ser mais simplificado. Por fim, em todos os testes de redistribuição realizados para as novas áreas foram obtidos resultados similares aos da área de eletricidade, ou seja, aumento do IBD com a redistribuição, e aprovações nos testes de hipótese de Mann-Whitney para todos os casos.

Sendo assim, embora o ideal fosse obter amostras com o tempo dispendido no exame de cada uma das áreas técnicas avaliadas para calibração específica do modelo, por meio das análises realizadas, e com a determinação dos critérios/condições decorrentes delas, foi possível garantir uma boa execução do método da ACP, um comportamento estatístico do IGC semelhante ao caso já validado com o tempo (com os dados da área de eletricidade) e o bom funcionamento dos enquadramentos em classes (funcionam bem para distribuição normal). Além disso, uma correlação similar com a principal variável de interesse (número de páginas de reivindicações) faz com que o modelo seja conservador, sendo que em conjunto com uma melhora do índice IBD, há uma confirmação de maior equilíbrio das redistribuições testadas. Portanto, considera-se que o modelo desenvolvido pode ser aplicado a outras áreas do INPI desde que seja realizada uma análise estatística prévia da amostra a ser utilizada e sejam atendidos os seguintes critérios mínimos: a) número mínimo de 20 amostras para simulação de casos com 3 variáveis; b) variância acumulada da componente principal $Y1$ elevada (cerca de 60% ou maior); c) amostra de dados do IGC_{Y1} com distribuição estatística aproximadamente normal; d) alta correlação dos dados de IGC_{Y1} com o número de páginas de

reivindicações (cerca de 0,90); e) valor alto de IBD nas simulações (cerca de 0,80 ou maior) sendo que, em caso de testes de redistribuição, verificar um aumento do IBD nas simulações.

Em suma, conforme discutido em detalhes ao longo da apresentação dos resultados, verificou-se que o modelo desenvolvido consegue representar de maneira bastante satisfatória o esforço/tempo de exame de pedidos de patente com aproveitamento de buscas de escritórios internacionais (tipo 6.21), sendo que para casos no limiar dos critérios de enquadramento adotados ocorrem poucas divergências e, nestes casos, as divergências ocorrem apenas nas faixas adjacentes. Ou seja, ocorrem divergências pontuais, não grosseiras e justificáveis dadas as limitações inerentes a um modelo e pesquisa deste tipo. Cumpre destacar que há dois fatores que possivelmente explicam as limitações e erros do modelo até o momento: i) erros inerentes ao modelo matemático, os quais ainda que em pequena quantidade, tendem a ocorrer dependendo das variáveis, amostras e critérios adotados; ii) erros ou desvios de medição no tempo por parte dos examinadores, o que poderia mover um enquadramento perto do limite das faixas.

Por fim, cabe destacar que foi desenvolvido um modelo funcional e de fácil aplicabilidade para avaliação e distribuição dos pedidos de patente do tipo 6.21 para os examinadores do INPI, que pode servir não somente como um embrião para a aplicação prática de uma primeira distribuição automática de pedidos aos examinadores, mas também servir de apoio à decisão dos gestores ao realizar a distribuição manualmente, contribuindo diretamente para a eficiência do próprio INPI.

Como sugestões de trabalhos futuros, há dois eixos principais que podem ser seguidos: i) ampliação do escopo de aplicação da metodologia desenvolvida; e ii) aperfeiçoamentos específicos do modelo matemático desenvolvido.

A primeira sugestão de trabalho é estender a aplicabilidade da metodologia desenvolvida para pedidos de patente sem aproveitamento de buscas realizadas por escritórios internacionais, ou seja, aqueles enquadrados como tipo 6.22 no plano de combate ao backlog. Além de ser utilizada para identificar possíveis adaptações da metodologia em função das características destes novos tipos de pedidos, tal expansão possibilitaria aplicá-la a praticamente todos os pedidos de primeiros exames a serem distribuídos pelo INPI, aprimorando ainda mais os seus processos e sua eficiência.

Uma segunda sugestão de trabalho, ainda com relação ao escopo de aplicação da metodologia, seria a realização de estudos com dados de outras áreas tecnológicas da DIRPA ainda não testadas, além de ampliação contínua dos testes para as áreas já realizadas.

Trabalhos seguindo essa linha possibilitarão identificar adaptações da metodologia em função de características específicas das áreas tecnológicas, além de sedimentar a sua confiabilidade e rendimento.

Outro trabalho interessante é a expansão da metodologia para aplicabilidade a outras etapas de exame (distribuição de segundos exames, etc.) ou mesmo à 2ª instância (distribuição de recursos). Trabalhos nesse sentido seriam importantes para adaptação da metodologia principalmente quanto às variáveis de interesse, dado que pedidos com primeiros exames já realizados devem considerar as manifestações e possíveis novos quadros reivindicatórios e, no caso de recursos, além da manifestação, toda a documentação dos exames realizados na 1ª instância.

Quanto aos aprimoramentos específicos do modelo matemático, poderiam ser realizados testes com a inclusão de novas variáveis e, além disso, testes com novas formas de enquadramento dos pedidos em classes como, por exemplo, a utilização de redes neurais ou outras técnicas de inteligência artificial. Tais trabalhos podem melhorar ainda mais a eficiência do modelo, na medida em que podem aumentar as correlações dos índices com o tempo dispendido de exame, e aumentar os acertos nos enquadramentos dos pedidos.

REFERÊNCIAS

ARCHONTOPOULOS, E., GUELLEC, D., STEVNSBORG, N., DE LA POTTERIE, B. V. P. & ZEEBROECK, N. V. (2007). **When small is beautiful: Measuring the evolution and consequences of the voluminosity of patent applications at the EPO.** Information Economics and Policy, 19(2), 103-132.

ARCHONTOPOULOS, E., VAN ZEEBROECK, N., STEVNSBORG, N., DE LA POTTERIE, B. V. P., & GUELLEC, D. **Patent inflation in Europe.** World Patent Information, 30(1), 43-52. 2008.

BEGNINI, S., ALMEIDA, L. E. D. F. **Grau de desenvolvimento dos municípios da mesorregião oeste catarinense: caracterização e classificação.** Revista Interações, v.17, n.4, p.547-560. 2016.

BRASIL. **Lei nº 9.279, de 14 de maio de 1996 [Lei da Propriedade Industrial].** Regula direitos e obrigações relativos à propriedade industrial. Disponível em <http://www.planalto.gov.br/ccivil_03/leis/L9279.htm> Acessado em Fevereiro, 2017.

CATTEL, R. B. (1966). **The Screen Test for the Number of Factors.** Multivariate Behavior Research, V. 1, pp. 245-276.

CESARE, R. & TIMOTHY, S. **Patent Examiner Specialization.** NBER Working Paper, 2018.

COCKNURN, I. M., KORTUM, S. & STERN, S. **Are All Patent Examiners Equal? The Impact of Examiner Characteristics.** NBER Working Paper 8980, 2002.

DELVIZIO, E. **Avaliação das estratégias da indústria para-petrolífera pelos respectivos depósitos de patentes.** In: Rio Oil & Gas Expo and Conference 2010, Rio de Janeiro, 2010.

DE SAINT-GEORGES, M. & DE LA POTTERIE, B. V. P. **A quality index for patent systems.** Research Policy, 42(3), 704-719. 2013.

FRIEL, C. M. **Notes on Factor Analysis.** Criminal Justice Centre, Sam Houston State University. 2009

GAMBARDELLA, A., HARHOFF, D. & VERSPAGEN, B. **The Value of European Patents**. *European Management Review*, 5(2), 69-84, 2008.

GIBBONS, J. D. **Nonparametric Statistical Inference**. New York: McGraw-Hill, 1971, 306 p.

GOLODNE, D. M., & SILVA, A. S. **Managing patent examiners' workload by using patent application volume data**. *Manuscripto submetido para publicação*, Junho, 2020.

HAIR, J. F. ET AL. **Análise multivariada de dados**. 6. ed. Porto Alegre: Bookman, 2009.

HARLOFF, D. & WAGNER, S. **Modeling the Duration of Patent Examination at European Patent Office**. GESY, Discussion Paper 170, 2006.

HERIG, L. C. T. **O Modelo das Componentes Principais: apresentação e interpretações**. Tese (mestrado) – Universidade Federal do Rio de Janeiro, Faculdade de Engenharia, 1979. 222p.

HONGYU, K., SANDANIELO, V. L. M., JUNIOR, G. J. de O. **Análise de componentes principais: resumo teórico, aplicação e interpretação**. *Engineering and Science*, V.1, 5ed. 2015.

HONGYU, K. **Comparação do GGE-biplot ponderado e AMMI-ponderado com outros modelos de interação genótipo × ambiente**. Tese (Doutorado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2015. 155p.

HOTELLING, H. **Analysis of a complex of statistical variables into principal components**. *The journal Educational Psychology*, Cambridge, v.24, p.498-520, 1933.

HOTELLING, H. **Simplified calculation of principal components**. *Psychometrika*, Williamsbourg, v.1, p.27-35, 1936.

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL – INPI. **Classificação Internacional de Patentes – CIP**. Disponível em <<http://ipc.inpi.gov.br/>> Acessado em Novembro, 2019.

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL – INPI. **INPI e EPO lançam parceria técnica e estratégica na área de patentes**. Disponível em <<http://www.inpi.gov.br/noticias/inpi-e-epo-lancam-parceria-tecnica-e-estrategica-na-area-de-patentes/>> Acessado em Março, 2020.

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL – INPI. **INPI tem novo Plano Estratégico com missão e visão 2021.** Disponível em <<http://www.inpi.gov.br/noticias/inpi-tem-novo-plano-estrategico-com-missao-e-visao-2021/>> Acessado em Março, 2020.

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL – INPI. **Instrução Normativa INPI nº 30**, 2013.

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL – INPI. **Instrução Normativa INPI nº 31**, 2013.

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL – INPI. **Novo INPI: Protocolo de Madri e Plano de Combate ao Backlog de Patentes.** Disponível em <<http://www.inpi.gov.br/noticias/protocolo-de-madri-e-plano-de-combate-ao-backlog-de-patentes/>> Acessado em Março, 2020.

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL – INPI. **Plano de Combate ao Backlog.** Disponível em <<https://www.gov.br/inpi/pt-br/servicos/patentes/plano-de-combate-ao-backlog>> Acessado em Agosto, 2019.

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL – INPI. **Procedimento de Avaliação de Subsídios de Terceiros ao Exame Técnico – DIRPA-P026**, 2015.

JOHNSON, R.A.; WICHERN, D.W. **Applied multivariate statistical analysis.** Madison: Prentice Hall International, 1998. 816p.

JOHNSON, R. A., WICHERN, D. W. (2002), **Applied Multivariate Statistical Analysis**, New Jersey: Prentice-Hall, Inc., 5 ed., 797p.

KAISER, H. F. **The varimax criterion for analytic rotation in fator analysis.** Psychometrika, v. 23, n. 3.p. 187-200, 1958.

KIM, Y. K., & OH, J. B. **Examination workloads, grant decision bias and examination quality of patent office.** Research Policy, 46(5), 1005-1019, 2017.

KUBRUSLY, L. S. **Um procedimento para calcular índices a partir de uma base de dados multivariados.** Pesquisa Operacional, Vol. 21, N^o1, p. 107-117, junho de 2001.

KUHN, J. M. & THOMPSON, N.C. **How to Measure and Draw Causal Inferences with Patent Scope**. Working Paper, 2017.

LEMLEY, M. A. & SAMPAT, B. **Examiner Characteristics and Patent Office Outcomes**. Review of Economics and Statistics 94(3), 817-827, 2012.

LOPES, L. F. D. **Análise de Componentes Principais Aplicada à Confiabilidade de Sistemas Complexos**. Tese de Doutorado. Universidade Federal de Santa Catarina, 2001. 138p.

MARCO, A. C., SARNOFF, J. D. & DEGRAZIA, C. **Patent Claims and Patent Scope**. Working paper. 2016.

MATRIZ CALCULATOR. **Cálculo dos Valores e Vectors Próprios**. Disponível em <<https://matrixcalc.org/pt/vectors.html>> Acessado em Setembro, 2020.

MEIRA, C. T.; PEREIRA, I. G.; FARAH, M. M.; PIRES, A. V.; GARCIA, D. A.; CRUZ, V. A. R. **Seleção de características morfofuncionais de cavalos da raça Manga larga Marchador por meio da análise de componentes principais**. Arquivo Brasileiro de Medicina Veterinária e Zootecnia, 65, 1843–1848. 2013.

NOETHER, G. E. **Introduction to Statistics: A Nonparametric Approach**. Boston: Houghton Mifflin, 1976, 292 p.

OLIVEIRA, L. G., SUSTER, R., PINTO, A. C., RIBEIRO, N. M. & SILVA, R. B. **Informação em Patentes: Ferramenta Indispensável para a Pesquisa e o Desenvolvimento Tecnológico**. Química Nova, Vol. 28, S36-S40, 2005.

ORGANIZAÇÃO MUNDIAL DA PROPRIEDADE INTELECTUAL – OMPI. **Classificação Internacional de Patentes – CIP**. Disponível em <<https://www.wipo.int/classifications/ipc/ipcpub>> Acessado em Novembro, 2019.

ORGANIZAÇÃO MUNDIAL DA PROPRIEDADE INTELECTUAL – OMPI. **Norma ST.9 - Manual de Informação e Documentação em Matéria de Propriedade Industrial**. 2013. Disponível em <<https://www.wipo.int/export/sites/www/standards/es/pdf/03-09-01.pdf>> Acessado em Novembro, 2019.

PAIVA, A. P. **Metodologia de Superfície de Resposta e Análise de Componentes Principais em Otimização de Processos de Manufatura com Múltiplas Respostas Correlacionadas**. Tese (Doutorado) – Universidade Federal de Itajubá, Instituto de Engenharia Mecânica, 2006. 257p.

PAIVA, A.L.C.; TEIXEIRA, R.B.; YAMAKI, M.; MENEZES, G.R.O.; Leite, C. .S.; TORRES, R.A. **Análise de componentes principais em características de produção de aves de postura.** Revista Brasileira de Zootecnia, 39, 285–288. 2010.

PALLANT, J. **SPSS Survival Manual.** Open University Press, 2007.

PEARSON, K. **On lines and planes of closest fit to systems of points in space.** Philosophical Magazine, 2(11), 559-572. 1901.

PHILIPP, M. **Patent filing and searching: Is deflation in quality the inevitable consequence of hyperinflation in quantity?** World Patent Information, 28(2), 117-121. 2006.

Portal Action. **Distribuição Normal.** Disponível em < <http://www.portalaction.com.br/probabilidades/62-distribuicao-normal>> Acessado em Dezembro, 2020.

Portal Action. **Teste de Kolmogorov-Smirnov.** Disponível em < <http://www.portalaction.com.br/inferencia/62-teste-de-kolmogorov-smirnov>> Acessado em Dezembro, 2020.

REGAZZI, A. J. **INF 766 - Análise multivariada.** Viçosa: Universidade Federal de Viçosa, Centro de Ciências Exatas e Tecnológicas. Departamento de Informática, 2001. 166p. Apostila de disciplina.

RENCHER, A. C. **Methods of Multivariate Analysis.** John Wiley and Sons, 2 ed., 740p. 2002.

SAVEGNAGO, R.P., CAETANO, S.L., RAMOS, S.B., NASCIMENTO, G.B., SCHMIDT, G.S., LEDUR, M.C. MUNARI, D.P. **Estimates of genetic parameters, and cluster and principal components analyses of breeding values related to egg production traits in a White Leghorn population.** Poultry Science, 90, p.2174-2188. 2011.

SANDANIELO, V. L. M. **Emprego de técnicas estatísticas na construção de índices de desenvolvimento sustentável aplicados a assentamentos rurais.** Tese (doutorado) - Universidade Estadual Paulista, Faculdade de Ciências Agrônômicas de Botucatu, 159p. 2008

SCREMIN, M. A. A. **Método para Seleção do Número de Componentes Principais com base na Lógica difusa.** Tese de Doutorado em Engenharia de Produção, Universidade Federal de Santa Catarina, 124p. 2003.

SENAC. **Programação VBA para Excel**. Apostila – Serviço Nacional de Aprendizagem Comercial, Porto Alegre – RS, 67p. 2010.

SILVA, M. C., SILVA, J. D. G., BORGES, E. F. **Análises de componentes principais para elaborar índices de desempenho no setor público**. Revista Brasileira de Biomedicina, v.33, n.3, p.291-309. 2015.

SOARES, PEDRO LEAL DE LIMA. **Aperfeiçoamento da Avaliação da Condição de Suficiência Descritiva nas Patentes**. Tese de Doutorado. Academia de Propriedade Intelectual do INPI, 186p. 2019.

SOARES, P. L. L., MOREIRA JR., C. V., RODRIGUES, R. C. & BRAGA, E. J. **O Papel do Termo Caracterizante nos Documentos de Patentes na Delimitação de Direitos em Propriedade Industrial**. Revista da ABPI – nº 161 – Jul/Ago 2019, 26-40. 2019.

SOCIAL SCIENCE STATISTICS. **Mann-Whitney Calculator**. Disponível em <<https://www.socscistatistics.com/tests/mannwhitney/default2.aspx>> Acessado em Março, 2021.

SOUZA, A. M. **Monitoração e Ajuste de Realimentação em Processos Produtivos Multivariados**. Tese de Doutorado em Engenharia de Produção, Universidade Federal de Santa Catarina, 166p. 2000.

VAN ZEEBROECK, N., DE LA POTTERIE, B. V. P. & GUELLEC, D. **Claiming more: the increased voluminosity of patent applications and its determinants**. Research Policy, 38(6), 1006-1020. 2009.

VICINI, LORENA. **Análise multivariada da teoria à prática**. Monografia (Especialização) – Universidade Federal de Santa Maria, Faculdade de Estatística, 2005. 215p.

VIALI, L. **Introduction to Statistics: A Nonparametric Approach**. Boston: Houghton Mifflin, 1976, 292 p.