

## Técnica Computacional Multi-Objetiva Para Definição de Políticas de Crédito Para Instituições de Microfinanças

**Autoria:** Eleonora Cruz Santos, Luiz Henrique Duczmal

### Resumo

Neste trabalho construímos um modelo computacional que possa dar suporte ao processo de concessão de crédito e gestão dos recursos financeiros das Instituições de Microfinanças (IMF). Criamos uma ferramenta de otimização multi-objetivo que busca maximizar simultaneamente o nível de eficiência e eficácia para a política de seleção de candidatos a crédito, com o objetivo de facilitar a sustentabilidade da IMF sem a dependência de doação de recursos. A principal novidade metodológica é a aplicação de uma técnica de Monte Carlo para testes de permutação a fim de gerar configurações sob a hipótese nula em problemas multi-objetivos, substituindo as técnicas econométricas clássicas, e as análises do tipo DEA e técnicas fuzzy, mais frequentemente encontradas em artigos computacionais da área. Utilizamos o novo conceito do vetor de significância, que permite realizar o cálculo de inferência da robustez dos resultados encontrados na fronteira ótima de Pareto.

### 1. Introdução

As Instituições de Microfinanças (IMF) brasileiras têm se constituído com apoio fundamental dos governos municipais e organismos internacionais e nacionais, donde se destaca a atuação do Banco InterAmericano de Desenvolvimento (BID), o BNDES, os bancos de desenvolvimento regional e as prefeituras municipais. A grande contribuição destas IMF tem se dado na inserção de cidadãos menos privilegiados na dinâmica da economia, bem como na redução da pobreza. Ademais, podem se constituir em possíveis instrumentos de políticas públicas e, por conseguinte, estarem inseridos dentro de um contexto maior de atuação do Estado, como por exemplo, na capacitação e treinamento de mão-de-obra factível de benefícios através do microcrédito. No entanto, a viabilização das IMF está intrinsecamente associada a sua capacidade de auto-sustentabilidade, o que, até o presente momento, tem sido um entrave para o seu crescimento e consolidação. Embora nos últimos anos o governo federal venha promovendo o microcrédito, seu enfoque também tem sido os bancos comerciais tradicionais, os quais, por sua vez, se distinguem em vários aspectos, sobretudo na possibilidade de receberem depósitos - instrumento fundamental na capacidade de expansão.

Neste trabalho buscamos construir um modelo (computacional) que possa dar suporte no processo de concessão de crédito e gestão dos recursos financeiros das IMF. A preocupação reside no fato de criarmos uma ferramenta que sinalize o maior nível de eficiência (medida como a seleção dos candidatos a crédito com menor probabilidade de inadimplência – elevados retornos) e, simultaneamente, o maior nível de eficácia (facilitar a sustentabilidade sem a dependência de doação de recursos; ou seja, promover elevado volume de recursos, ou do tamanho da carteira) para a política de seleção de candidatos a crédito. Diante do exposto, nosso objetivo inicial é construir a fronteira de Pareto com base na maximização dos dois objetivos (eficiência e eficácia) para as políticas de crédito factíveis no espaço objetivo. Em seguida, buscamos verificar se os resultados contidos na fronteira de Pareto representam, ou não, políticas significativamente melhores em termos de eficiência e eficácia. Caso tais políticas sejam significativamente melhores, discutimos critérios quantitativos para distinguir qual é a melhor dentre as possíveis soluções apontadas na fronteira de Pareto. A estatística de teste utilizada para encontrar a melhor política da fronteira é denominada vetor de significância. O trabalho utiliza simulações numéricas de Monte Carlo, computacionalmente intensivas, e por esta razão, optamos pela programação em linguagem C.

## 2. Definição do Problema e dos Multi-Objetivos para Otimização

Os conceitos de eficiência e eficácia têm ampla aplicação nas ciências sociais aplicadas e exatas. Naturalmente que são vastas suas definições e, dada a natureza deste trabalho, busca-se levantar na literatura os conceitos de eficiência e eficácia mais próximos à área de finanças. No entanto, cabe mencionar que, na área de estudos organizacionais, mais especificamente na discussão entre centralização e descentralização, há trabalhos como os de Christensen e Knudsen (2002) e Sah e Stiglitz (1986) que buscam construir uma fronteira eficiente do desenho das organizações e definem, de formas distintas, o conceito subjacente a este tipo de eficiência. Sah e Stiglitz (1986) utilizam as definições de organização hierárquica e “poliárquica”, cuja diferença reside no fato de, na primeira, uma decisão precisar do consenso de todos os membros para se aceitar um projeto e, na segunda esta decisão poder caber a um único membro. Christensen e Knudsen (2002) incluem formas intermediárias de hierarquia, incluindo um número arbitrário de agentes para avaliar a qualidade dos projetos recebidos. A construção da fronteira do desenho organizacional. A organização decide de acordo com a movimentação do agente.

Especificamente na literatura econômica voltada para finanças, ainda é escassa a produção de trabalhos lançando mão destas técnicas, donde se destaca Chang et al (2000), Shoaf e Foster (1996), Ehrgott et al (2004), Ruspini e Zwir (1999), Mullei e Beling (1998) e Mukerjee et al (2002). Destes trabalhos, os três primeiros fazem variações sobre o modelo de média-variância de Markowitz e utilizam da técnica de algoritmo genético para estimar os resultados ótimos para os problemas de otimização de carteiras.

Também fazendo uma analogia ao modelo de média-variância de Markowitz (1959), nosso trabalho busca otimizar dois objetivos fundamentais na garantia de auto-sustentabilidade das IMF: a *eficiência*, definida como uma medida da qualidade deste crédito, por busca mensurar, do total de crédito concedido, quanto efetivamente retornou ao sistema no prazo definido em contrato. A *eficácia*, definida como uma medida de volume de crédito, associada ao volume de crédito concedido (volume da carteira ativa) e retornado à IMF no período em análise. Formalmente tem-se que (considerado um período fixo de 12 meses):

$$ef_{i,j} = \sum_{i=1}^{12} \sum_{j=1}^M \frac{vr_{i,j} * fd_{i,j}}{vp_{i,j}}$$

onde:

$ef_{i,j}$  é a eficiência das  $i$  parcelas dos  $j$  clientes (inclui os candidatos recusados)

$vr_{i,j}$  é o valor da  $i$ -ésima parcela do  $j$ -ésimo cliente que foi paga

$vp_{i,j}$  é o valor da  $i$ -ésima parcela contratada a ser paga pelo  $j$ -ésimo cliente

$fd_{i,j}$  é o fator de desconto da  $i$ -ésima parcela paga do  $j$ -ésimo cliente

$M$  é o número de clientes

O fator de desconto foi construído com base na Resolução nº 2.682 do Banco Central do Brasil, que define os critérios de classificação das operações de crédito e regras para a constituição de provisão para créditos de liquidação duvidosa. Nesta Resolução há 9 (nove) níveis de classificação das operações de crédito, correspondente ao número de dias de atraso de cada parcela. Com base nesta classificação, construímos uma matriz de fatores de desconto para as parcelas, na qual, para cada parcela em atraso, desconta-se a TAC (taxa de abertura de crédito) e a taxa de juros “carregada” do início do contrato até a referida parcela.

A medida construída de eficiência estaria penalizando os pagamentos efetuados em atraso e permitindo que se mensure, em uma escala de 0 a 1, a pontualidade dos clientes da Instituição. É neste sentido que entendemos a eficiência como uma medida de qualidade dos clientes que compõem a carteira de crédito da IMF. Para a eficácia, a definição formal é:

$$ec_{i,j} = \sum_{i=1}^{12} \sum_{j=1}^M vg_{i,j}$$

onde:

$ec_{i,j}$  é a eficácia das  $i$  parcelas dos  $j$  clientes (inclui os candidatos recusados)

$vg_{i,j}$  é o valor pago da  $i$ -ésima parcela do  $j$ -ésimo cliente

A eficácia nos dá a magnitude do crédito concedido que retornou ao sistema, independente do tempo de retorno, ou seja, é uma medida mais voltada à quantidade, e embute os juros e mora das parcelas pagas em atraso. Nos dá também uma noção do impacto do ingresso de novos clientes sobre a carteira de crédito.

Após definirmos os objetivos, podemos formalizar nosso problema multi-objetivo como:

$$\max_t (ef, ec)$$

$$\text{sujeito a } \begin{cases} p \\ c \\ ge \\ g \\ r \\ fc \\ vc \end{cases}$$

onde:

$ef$  é a medida de eficiência

$ec$  é a medida de eficácia

$t$  é o tempo, dividido em dois períodos de 12 meses cada;

$p$  é o posto de atendimento bancário;

$c$  é o estado civil, decomposto em casados, solteiros e outros;

$ge$  é o grau de escolaridade, decomposto em 1º. grau, 2º. grau e superior (completos ou não)

$g$  é o gênero, podendo ser homem e mulher, ou cada um separadamente;

$r$  é a renda líquida disponível do indivíduo, descontadas as despesas fixas mensais

$fc$  é a finalidade do crédito, decomposto em capital fixo, de giro e misto;

$vc$  é o valor do crédito, definido em três categorias: até R\$999,00, até R\$1.999,00 e R\$2.000,00 e mais.

Todas as variáveis foram contempladas pelas combinações factíveis entre suas respectivas categorias.<sup>i</sup>

Em síntese, neste trabalho abordamos o problema de maneira multi-objetiva e buscamos otimizar os dois objetivos, eficiência e eficácia, simultaneamente. Para isso implementamos um novo algoritmo de simulação computacional de Monte Carlo, com o intuito de avaliarmos a significância de regras hipotéticas que regeriam a concessão de crédito em uma IMF. Nossa contribuição final é a escolha de uma política creditícia<sup>ii</sup> que possa contribuir, de maneira ótima, aos requisitos de auto-sustentabilidade. Desenvolvemos um estudo de caso utilizando dados reais de uma IMF brasileira, denominada Associação de Crédito Popular, mais conhecida como Banco do Povo®, cuja sede se situa na cidade de Belo Horizonte.

### 3. A Base de Dados

A base de dados utilizada inclui todos os clientes pessoas físicas que tiveram seu crédito aprovado e iniciaram seu processo de empréstimo junto à Instituição, do período de setembro de 2004 a agosto de 2006, para um único posto de atendimento bancário (ou agência) da Instituição. Foram coletadas todas as informações sócio-econômicas, bem como o histórico de

pagamento das parcelas de crédito para cada contrato individualmente. Trabalhou-se apenas com os dados referentes ao PAB da cidade de Belo Horizonte, por ser o único com informações sócio-econômicas disponíveis dos candidatos solicitantes que tiveram o crédito recusado. Ao todo trabalhamos com uma base de 543 contratos de crédito e selecionamos, aleatoriamente, mais 210 candidatos a crédito que tiveram sua solicitação indeferida. Enfim, a base final de dados para as simulações contou com 753 contratos de crédito distribuídos em dois períodos de 12 meses cada.

#### 4. Metodologia

Esta seção descreve (1) *a técnica da permutação cíclica*, utilizada neste trabalho para se gerar aleatoriedade aos dados e testar a hipótese nula subjacente ao modelo de seleção de candidatos a crédito para instituições de microfinanças, bem como (2) *os vetores de significância*, construídos para se avaliar a robustez dos resultados encontrados.

##### 4.1. Teste de Permutação Cíclica

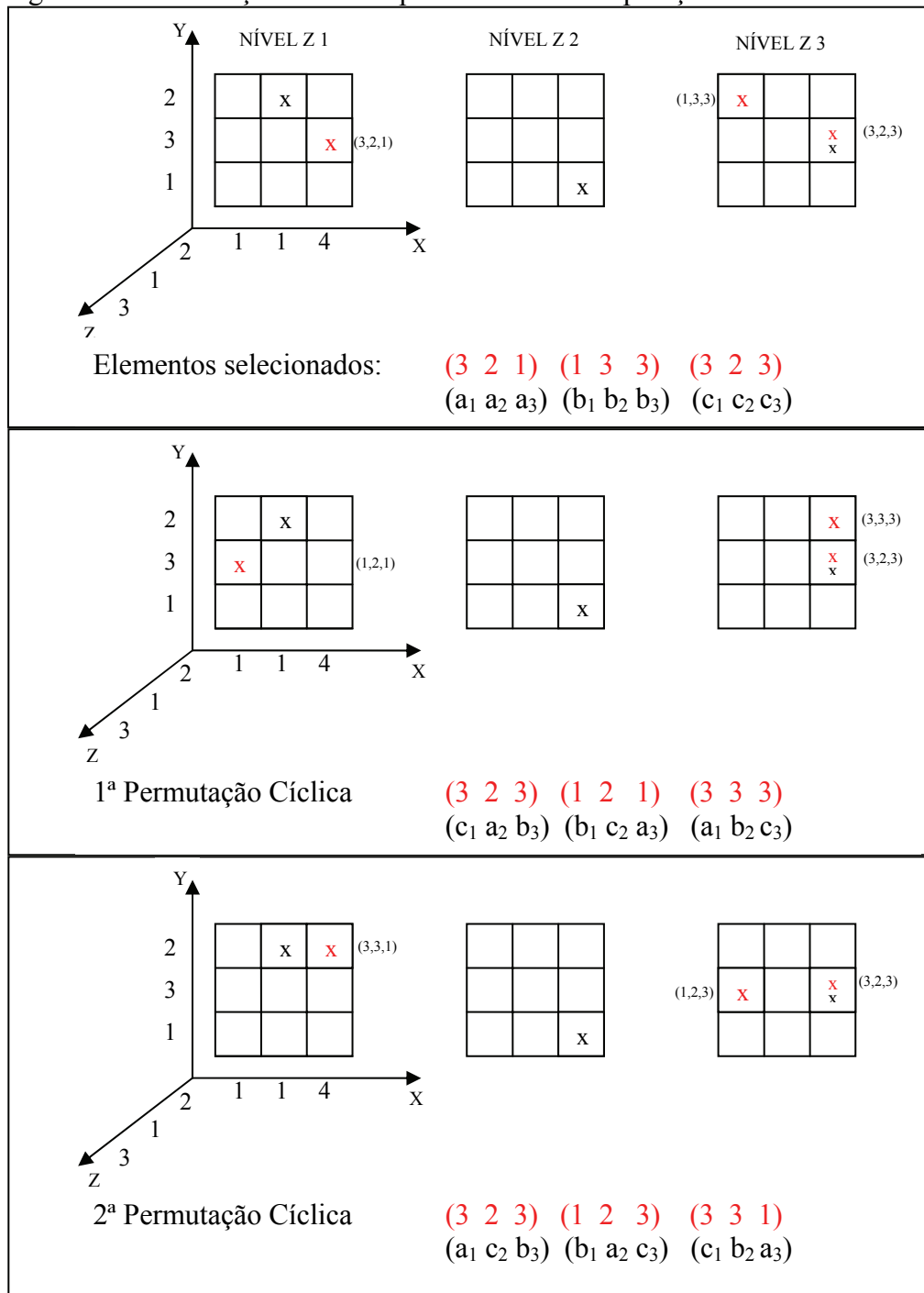
Manly (1997) define um *teste de permutação* como um modelo de probabilidade para se estimar a significância estatística das soluções ótimas obtidas com base nos verdadeiros dados da população em análise. O teste de permutação pode ser visto como uma ferramenta que permite se gerar aleatoriedade aos dados observados disponíveis sobre um determinado fenômeno que se queira investigar. No caso específico deste trabalho, cujo objetivo primeiro é construir um modelo de política de crédito com base nas variáveis de seleção dos candidatos, bem como em seus respectivos históricos de crédito, as soluções aleatorizadas são calculadas sobre um subconjunto de todas as possibilidades de arranjo entre as variáveis de seleção dos candidatos, sob as quais se têm os referidos históricos de crédito.

O *teste de permutação cíclica* é um tipo de teste de permutação que permite obter, a partir da distribuição multivariada para os dados observados, uma nova distribuição que elimine a dependência possivelmente existente entre as variáveis, mas que mantenham intactas as distribuições marginais, através do *Método de Monte Carlo*. A significância das soluções calculadas é extraída através da comparação dos resultados gerados *vis-à-vis* aqueles observados originalmente.

A FIG.1 ilustra detalhadamente os passos desta técnica. Suponha uma população de seis indivíduos num espaço tridimensional, cujos indivíduos (3, 2, 1), (1, 3, 3) e (3, 2, 3) sejam escolhidos aleatoriamente. Nesta figura, são ilustradas as duas possíveis permutações cíclicas entre os indivíduos. Após a realização das permutações, é fácil perceber que não houve alteração das frequências relativas das variáveis selecionadas, garantindo-se o pressuposto de manutenção das marginais. Em ambas as possíveis permutações, as marginais não se alteraram, uma vez que a mudança é feita par-a-par, dentro de um mesmo espaço físico. Essa propriedade é de fundamental importância, pois permite a criação de uma população modificada que conserva as distribuições marginais da população original.

Ademais, a possível dependência entre as variáveis no conjunto de indivíduos dos dados observados foi desfeita, gerando-se aleatoriedade aos dados originais e permitindo-se, a partir daí, que se teste os resultados originais em relação àqueles gerados aleatoriamente. Em outras palavras, *esta técnica permite que se gere configurações sob a hipótese nula de que não existe dependências entre as variáveis, a qual, por sua vez, busca testar se a configuração original se destaca em relação a outras configurações permutadas.*

Figura 1b - Permutação Cíclica Aplicada em uma População de Seis Indivíduos



Fonte: elaborado pelos autores

#### 4.2. A Permutação Aplicada ao Modelo de Seleção de Política de Crédito

Como dito no início da sub-seção 4.1, a permutação cíclica, ao gerar aleatoriedade aos dados observados disponíveis sobre um determinado fenômeno que se queira investigar, permite-nos, em última instância, que estimemos a significância das soluções calculadas em relação aos verdadeiros dados da população em análise. No caso específico deste trabalho, cujo objetivo primeiro é construir um modelo de política de crédito com base nas variáveis de seleção dos candidatos, bem como em seus respectivos históricos de crédito, as soluções calculadas são todas as possibilidades de arranjo entre as variáveis de seleção dos candidatos, sob as quais se têm os referidos históricos de crédito.

A significância das soluções calculadas é extraída através da comparação dos resultados gerados *vis-à-vis* aqueles observados originalmente. E, para este trabalho, a hipótese a ser testada pode ser assim definida:

$H_0$  : a política original *não se encontra* na Fronteira de Pareto em termos de eficiência e eficácia da seleção dos candidatos a crédito;

$H_1$  : a política original *se encontra* na Fronteira de Pareto em termos de eficiência e eficácia da seleção dos candidatos a crédito.

Aplicamos a técnica de permutação cíclica para a seleção da(s) política(s) de crédito sobre 7 (sete) variáveis<sup>iii</sup> de seleção de crédito que se mostraram mais relevantes, a constar: (i) posto de atendimento bancário, (ii) estado civil, (iii) gênero, (iv) grau de escolaridade, (v) renda líquida disponível, (vi) finalidade do crédito, (vii) valor do crédito. Cada uma das sete variáveis gerou possibilidades distintas de resultados. Por exemplo, a política de seleção de crédito poderia ser aquela que contemplasse: indivíduos, homens e mulheres, casados, com até o primeiro grau de escolaridade (completo ou não), renda líquida disponível inferior a R\$500,00 (quinhentos reais), com finalidade de crédito voltada para capital fixo, e cujo valor do crédito fosse de até R\$960,00 (novecentos e sessenta reais). Dentre as variáveis relevantes, acima definidas, selecionou-se possíveis sub-categorias, dentre cada uma delas, que fossem factíveis para a concessão do crédito. As variáveis e suas sub-categorias foram assim definidas:

- estado civil: (i) solteiros e casados; (ii) solteiros, casados e outros.
- gênero: (i) homens; (ii) mulheres; (iii) homens e mulheres.
- grau de escolaridade: (i) até 1º grau completo; (ii) até 2º grau completo; (iii) até superior completo.
- renda líquida disponível (a renda descontada todos os gastos mensais do cliente; definida a partir do questionário sócio-econômico aplicado pelo agente de crédito no domicílio ou estabelecimento comercial onde reside ou trabalha, respectivamente, o candidato ao crédito): (i) até R\$499,99 ; (ii) acima de R\$499,99. Inicialmente utilizou-se as duas categorias, mas por não se mostrarem significativas, foram agrupadas em uma única categoria de análise.
- finalidade do crédito: (i) crédito para capital fixo e capital de giro; (ii) crédito para capital fixo, giro e misto (este último definido como a composição de capital fixo e de giro) conjuntamente.
- valor do crédito: (i) até R\$999,99 (novecentos e noventa e nove reais e noventa e nove centavos); (ii) até R\$1.999,99 (hum mil, novecentos e noventa e nove reais e noventa e nove centavos); (iii) acima de R\$1.999,99 (hum mil, novecentos e noventa e nove reais e noventa e nove centavos).

Ao todo, temos 7 variáveis, das quais PAB (posto de atendimento bancário) e renda líquida só contemplam uma única sub-categoria.<sup>iv</sup>

### 4.3. Os Vetores de Significância

A construção de uma hipótese a ser testada pressupõe que se aplique(m) algum(ns) teste(s) de validação para se aceitar ou não a referida hipótese. Seja qual for a técnica que se esteja utilizando, a idéia básica consiste em construir a hipótese, testá-la e rejeitá-la ou não, com um determinado grau de significância. Nos modelos econométricos, por exemplo, supondo-se que os dados seguem uma distribuição Normal, as estatísticas  $t$  e  $F$  informam, a um determinado nível de significância, se  $b$  é um estimador não-viesado de  $\beta$ , não rejeitando, assim, a hipótese nula,  $H_0$ .

Neste trabalho, a hipótese nula,  $H_0$ , não vai ser testada utilizando-se modelos econométricos, uma vez que o tratamento e análise dos dados segue uma abordagem essencialmente “empírica-descritiva”. Por conseguinte, optou-se pela criação de uma estatística de teste nos moldes da estatística  $t$ .

Dados  $R$  valores amostrais  $b_i, i = 1, \dots, R$  que são instâncias de uma variável aleatória  $\beta$ , e um valor dado  $\mu_0$ , a estatística  $t$  com  $R-1$  graus de liberdade é definida como

$$t = \frac{b - \mu_0}{\sigma_b / \sqrt{R}},$$

onde:

$b$  é a média amostral

$\sigma_b$  é o desvio-padrão amostral, definido como

$$\sigma_b = \sqrt{\left( R \left( \sum_{i=1}^R b_i^2 \right) - \left( \sum_{i=1}^R b_i \right)^2 \right) / R(R-1)}$$

A medida construída para aferirmos a significância da hipótese a ser testada é denominada Vetor de Significância, formado por dois componentes, como se segue:

$$dp_{x_0, ef_{x_0}} = \frac{R * ef_{x_0} - \sum_{i=1}^R ef_{x_i}}{\sigma_{ef} * \sqrt{R}} \quad \text{e} \quad dp_{x_0, ec_{x_0}} = \frac{R * ec_{x_0} - \sum_{i=1}^R ec_{x_i}}{\sigma_{ec} * \sqrt{R}}$$

onde:

$dp_{x_0, ef_{x_0}}$  é a diferença entre os resultados de eficiência da política avaliada e a média das permutadas, em número de desvios-padrões das políticas permutadas;

$dp_{x_0, ec_{x_0}}$  é a diferença entre os resultados de eficácia da política avaliada e a média das permutadas, em número de desvios-padrões das políticas permutadas;

$ef_{x_0}$  e  $ef_{x_i}$  são os resultados de eficiência para a política avaliada,  $x_0$ , e as políticas permutadas  $x_i$ , respectivamente;

$ec_{x_0}$  e  $ec_{x_i}$  são os resultados de eficácia para a política avaliada,  $x_0$ , e as políticas permutadas  $x_i$ , respectivamente;

$\sigma_{ef}$  e  $\sigma_{ec}$  são os desvios-padrões do resultado das políticas permutadas  $x_i$  para eficiência e eficácia, respectivamente;

Para encontrarmos o resultado conjunto de eficiência e eficácia, extrai-se a raiz quadrada da soma dos quadrados dos “vetores parciais de significância” de eficiência e eficácia da política avaliada em relação às permutadas. Por fim, o *Vetor de Significância* é definido como:

$$vS_{x_0} = (dp_{x_0, ec_{x_0}})^{\hat{i}} + (dp_{x_0, ef_{x_0}})^{\hat{j}}$$

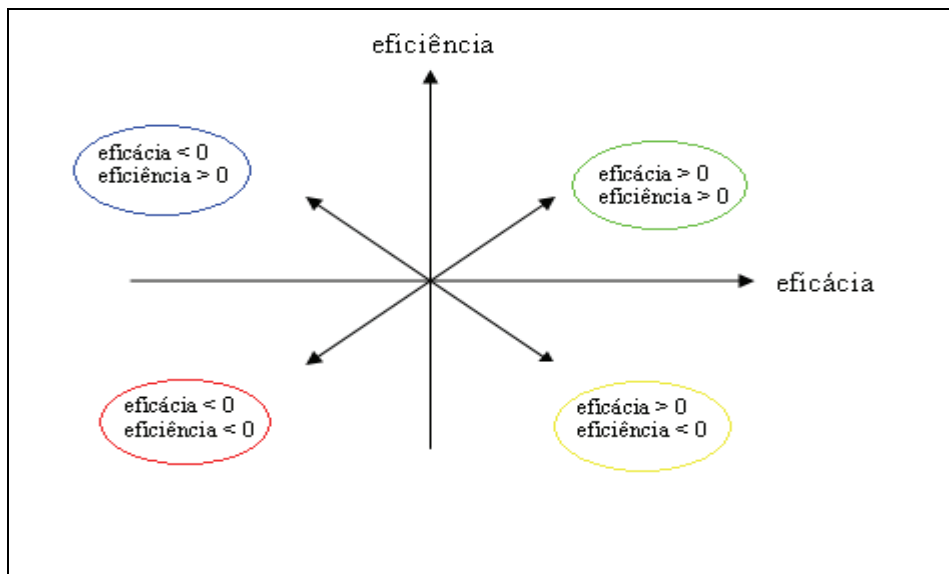
onde:  $\hat{i}$  e  $\hat{j}$  são os vetores unitários de eficácia e eficiência, respectivamente.

O módulo do vetor de significância é definido como:

$$\|vS_{x_0}\| = \left( (dp_{x_0, ec_{x_0}})^2 + (dp_{x_0, ef_{x_0}})^2 \right)^{\frac{1}{2}}$$

indicando o quão afastada a política original (não permutada) está, em termos de eficiência e eficácia, de políticas simuladas que sejam inteiramente aleatórias. O vetor de significância busca mostrar basicamente qual o ganho (ou não), em termos de desvios-padrões, de se adotar uma determinada política (factível) que faz parte do espaço em relação a outras quaisquer. Em outras palavras, mede a relevância da política definida. Este mesmo procedimento de cálculo é aplicado à base expandida, ou seja, àquela que inclui os candidatos recusados. Na verdade, é com esta base que buscamos avaliar os resultados multi-objetivos para eficiência e eficácia.

**Figura 2 – Representação dos Vetores de Significância (cor e direção)**



Fonte: elaborado pelos autores

Podemos notar que há quatro possibilidades de resultados de pares ordenados de eficiência e eficácia gerados pelo Vetor de Significância, como ilustra a FIG.2 abaixo, conforme a direção em cada um dos quatro quadrantes do plano cartesiano (eficiência *versus* eficácia).

O Vetor de Significância, por contemplar o resultado de uma determinada política  $x_0$  em número de desvios-padrões do conjunto de políticas  $x_i$ , permite aferir, também, com quantos desvios-padrões em relação a tais políticas,  $x_0$  é (visualmente) significativa.

Nos resultados apresentados na próxima seção, os vetores de significância apresentam tanto as cores circuladas (verde, vermelha, azul e amarela) quanto a direção dos resultados dos pares ordenados de eficiência e eficácia da FIG.2, sendo verde quando ambos os objetivos são positivos, amarelos quando só a eficácia é positiva, vermelho quando ambos são negativos e azul quando só eficiência é positiva.

## 5. O Modelo Multi-Objetivo de Política de Crédito

O modelo computacional de definição de políticas de crédito para IMF foi construído seguindo uma série de etapas, cada qual gerando um ou mais algoritmos que iriam compor o programa final. Estas etapas podem ser sumarizadas como se segue:

- primeiramente, definimos uma série de critérios para se estimar o histórico de crédito dos candidatos que foram recusados, mas que passaram a fazer parte da base de dados utilizada pelo modelo computacional;
- segundo, criamos o algoritmo que gera o histórico de cada candidato recusado, sendo este último extraído aleatoriamente da base de dados dos candidatos recusados da Instituição;



- em terceiro, construímos os algoritmos que mensuraram eficiência e eficácia com base em suas definições conceituais definidas na seção 2;
- em quarto, desenvolvemos o algoritmo que executa a permutação cíclica;
- em quinto, após se encontrar o par ordenado de eficiência e eficácia para cada política de crédito gerada com e sem permutação cíclica, construímos o algoritmo que mostrava graficamente as camadas de Pareto, ou seja, o par ordenado dos resultados de eficiência e eficácia para cada resultado encontrado de política de crédito com e sem permutação;
- e, por fim, desenvolvemos o algoritmo que permitiu visualizar a significância das políticas encontradas nas camadas de Pareto;

### 5.1 A Construção do Histórico dos Solicitantes Recusados

No contexto deste trabalho, como os dados observados dizem respeito somente àqueles candidatos a crédito que tiveram sua solicitação aceita, faz-se necessário a inclusão dos candidatos cujas solicitações foram indeferidas, para que se componha todo o possível universo de candidatos a crédito do posto de atendimento bancário analisado. De posse deste universo, a permutação cíclica passa a ser corretamente aplicada. Para tanto, faz-se necessário construir o histórico de pagamento do crédito para candidatos recusados. O passo inicial é buscarmos, na base de dados da instituição em análise, os dados sócio-econômicos dos possíveis candidatos que tiveram crédito indeferido. A seleção se deu por amostragem, para o mesmo período em que os dados dos candidatos aceitos foram analisados. Como os candidatos recusados não possuem históricos de crédito, construímos um histórico de pagamento das parcelas de crédito partindo do pressuposto de que tal histórico segue uma distribuição de Poisson não-homogênea. Duas variáveis são fundamentais na construção deste histórico de pagamento: a primeira é a parcela de crédito, e a segunda é a diferença entre a data de vencimento e a data de pagamento da parcela de crédito, e denominada como número de dias de atraso. A variável número de dias de atraso, por não ter um comportamento homogêneo ao longo do processo de pagamento das parcelas de um contrato, ou em termos genéricos, por apresentar uma taxa de ocorrência que pode mudar ao longo do período, foi tratada como seguindo uma distribuição de Poisson não-homogênea.<sup>v</sup> Este pressuposto é fundamental para o cálculo do número de dias de atraso das parcelas dos contratos dos candidatos recusados, com base no comportamento dos clientes. A etapa de construção do histórico de pagamento dos candidatos recusados consistiu, basicamente, em definir vários critérios, a constar:

1. assumiu-se como parâmetro de valor de crédito para cálculo do valor das parcelas, o montante solicitado de acordo com a pesquisa sócio-econômica realizada pela instituição junto ao candidato solicitante;
2. definiu-se como 12 meses o número de parcelas por ser este valor a moda do número de parcelas verificada nas estatísticas descritivas (ver sub-seção 3.2);
3. calculou-se o número de parcelas em função dos critérios adotados pela instituição solicitante. Em linhas gerais o crédito solicitado segue um cálculo padrão, havendo somente duas taxas de juros aplicadas. Para contratos com crédito inferior a R\$960,00 (novecentos e sessenta reais), a taxa de juros é de 2,00% a.m., e para valores iguais ou superiores a R\$960,00 (novecentos e sessenta reais), a taxa de juros aplicada é de 3,90% a.m.. Formalmente tem-se que:

$$vp = vc * \left[ \left( 1 + \frac{tac}{100} \right) * (1 + ff) \right]$$

onde:  $vp$  = valor da parcela

$vc$  = valor do crédito solicitado

$tac$  = taxa de abertura de crédito

$fj$  = fator de juros – conversão da taxa de juros adotada em fator pela tabela Price. Para a taxa de juros de 2,00% a.m., o fator da tabela Price é de 0,09456 e para a taxa de juros de 3,90% a.m., o fator aplicado é de 0,1059.

4. parcelas em atraso sofreram correção de mora e multa. O valor da mora definida pela instituição é de 0,33% a.d. e da multa é de 2,00%, ambos incidentes sobre o valor da parcela em atraso. Para o cálculo do valor de mora, é necessário que se tenha o número de dias de atraso, a qual é conceitualmente definida a seguir.

O número de dias de atraso dos clientes da Instituição é uma informação relevante para a definição do número de dias de atraso para os candidatos recusados. O histórico de dias de atraso do pagamento das parcelas de crédito dos solicitantes recusados foi construído com base na proximidade das características destes indivíduos em relação aos clientes pertencentes à base de dados da instituição. Melhor explicando, buscou-se a diferença entre os valores observados das características dos candidatos aceitos (clientes) e as mesmas características dos candidatos recusados. Em termos práticos, construímos uma matriz das diferenças das características entre cada candidato recusado e todos os clientes, e a partir daí ordenamos, em ordem crescente, os resultados para cada candidato recusado e selecionamos os 10 resultados menores, ou seja, aqueles 10 clientes que mais se aproximavam dos candidatos recusados. Formalmente tem-se que:

$$MD_{i,j} = \sum_{k=1}^6 |MS_{[i,k]} - MC_{[j,k]}|$$

onde:  $MD_{i,j}$  é a matriz das diferenças entre as características dos candidatos recusados e os clientes

$MS_{i,k}$  é a matriz das características dos candidatos recusados

$MC_{j,k}$  é a matriz das características dos clientes

A matriz  $MD'_{j,i}$  é a transposta de  $MD_{i,j}$ , e a ordenação, das diferenças de cada candidato recusado em relação aos demais clientes, em ordem decrescente, permite-nos identificar quais os clientes mais próximos a cada candidato recusado. Assim, valores mais próximos de zero referem-se aos clientes que apresentam características mais próximas àqueles solicitantes que tiveram seu crédito indeferido.

Da matriz  $MD'_{j,i}$  selecionamos os dez primeiros valores, de cada vetor de diferenças entre o candidato solicitante e os clientes da base de dados, relacionando-os a seus respectivos clientes e históricos de crédito. Assim, para cada candidato recusado identificamos os 10 clientes<sup>vi</sup> mais próximos no que diz respeito às características para seleção do candidato ao crédito.

O número de dias de atraso dos candidatos recusados, definida como uma variável que segue uma distribuição de Poisson não-homogênea, foi calculado da seguinte forma:

- 1 extrairmos o histórico do número de dias de atraso dos 10 clientes mais próximos ao solicitante;
- 2 ponderamos cada histórico através de pesos gerados aleatoriamente em um intervalo de 0 a 1;
- 3 geramos, aleatoriamente<sup>vii</sup>, o início de cada pagamento;
- 4 selecionamos, aleatoriamente, quais candidatos recusados teriam passado por algum processo de renegociação.

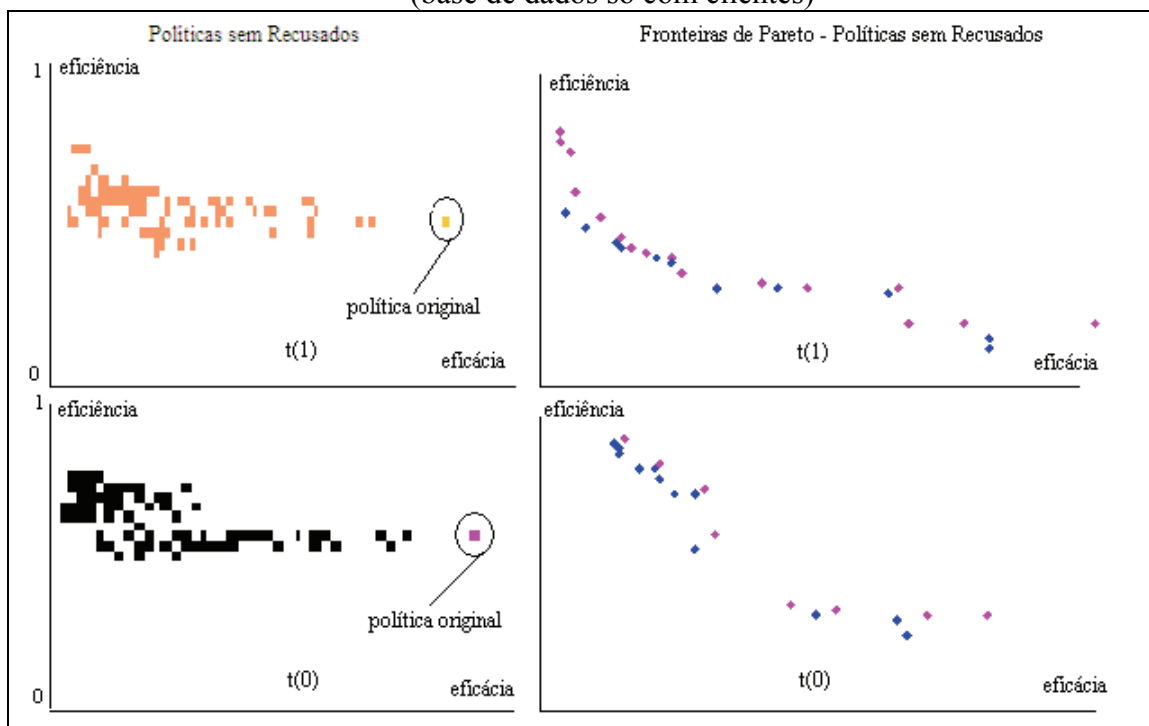
Para o cálculo da média ponderada de cada histórico de crédito, geramos 9 números aleatórios para dar pesos distintos para cada histórico dos clientes vizinhos próximos, e assim permitir

maior aleatoriedade aos resultados de cada solicitante. Estes números foram ordenados e posteriormente extraímos a diferença entre os mesmos, construindo, por fim, o valor de cada peso. A cada contrato aplicamos um dos 10 pesos gerados ao final deste processo.

### 5.2 Aplicação da Permutação Cíclica com e sem Candidatos Recusados

A FIG.3 e 4 apresentam os resultados da construção da Fronteira de Pareto com dois níveis, sendo o primeiro nível definido na cor rosa, e o segundo na cor azul. A razão para se escolher dois níveis é simples: dado tanto a proximidade entre ambos os níveis, quanto o fato de todos os pontos ali representados serem ótimos de Pareto, trabalhar com os dois níveis estaria se reduzindo o risco de exclusão de pontos que poderiam ser tratados como de fronteira.

Figura 3 – Resultados de Eficiência e Eficácia e Respectivas Fronteiras de Pareto (base de dados só com clientes)

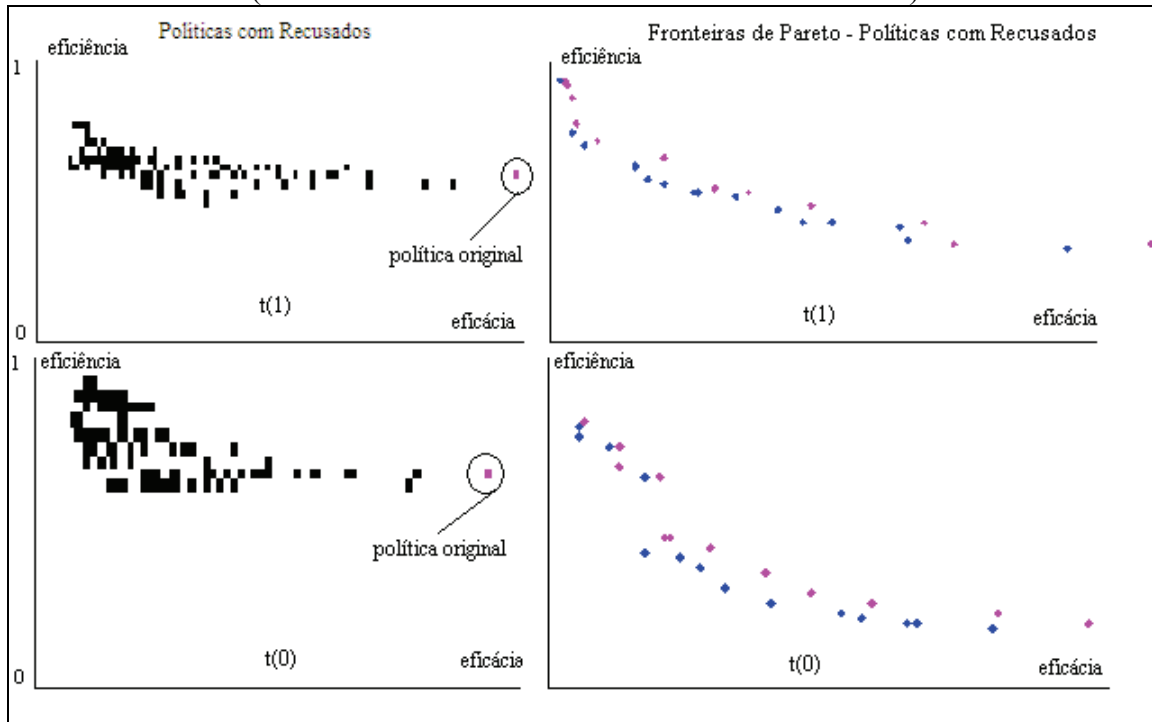


Fonte: elaborado pelos autores com dados do Banco do Povo

As políticas que incluem os resultados só para a base dos clientes e suas respectivas fronteiras estão conjuntamente representadas na FIG.3, ao passo que a FIG.4 inclui também os resultados dos candidatos recusados. Como pode se notar na FIG.3, o argumento de se selecionar dois níveis de não-dominância é, neste caso, bastante procedente, haja vista a proximidade dos resultados verificados em algumas políticas. Já na FIG.4 notamos que, em grande parte dos resultados que incluem os recusados, claramente há um distanciamento maior entre o nível 1 e 2.

Destes resultados surgem duas questões: primeira, quais são as políticas que representam o ótimo de Pareto(?), e segunda, dentre estas políticas, qual(is) atende(m) simultaneamente eficiência e eficácia ao maior nível de significância(?).

Figura 4 – Resultados de Eficiência e Eficácia e Respectivas Fronteiras de Pareto (base de dados com clientes e candidatos recusados)

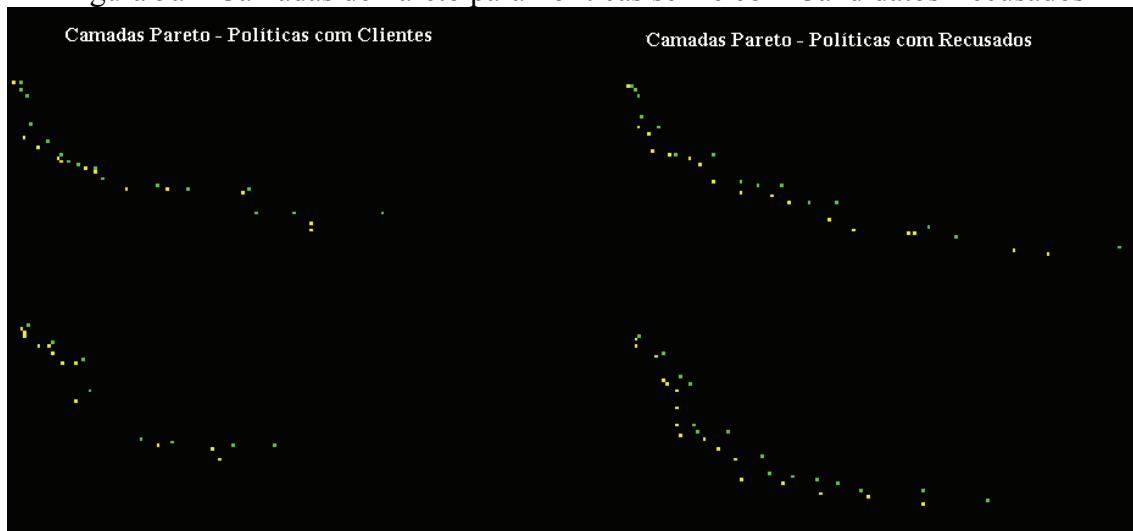


Fonte: elaborado pelos autores com dados do Banco do Povo

### 5.3 Os Campos dos Vetores de Significância e sua Interpretação

Os vetores que interessam à análise são aqueles que compõem a Fronteira de Pareto (FIG.5a), e estão representados na FIG.5b. Na Fig.5a, os pontos verdes representam a Fronteira de Pareto e os amarelos formam o segundo nível de não-dominância. Dentre os pontos de ótimo, aqueles cujos vetores de significância também estão representados pela cor verde são a resposta para o problema multi-objetivo que está se investigando, pois indicam ganhos de eficiência e eficácia em relação às demais políticas. Ademais, aqueles com maior número de desvios-padrões são também os que atendem, a um maior nível de significância, os dois objetivos simultaneamente.

Figura 5a – Camadas de Pareto para Políticas sem e com Candidatos Recusados

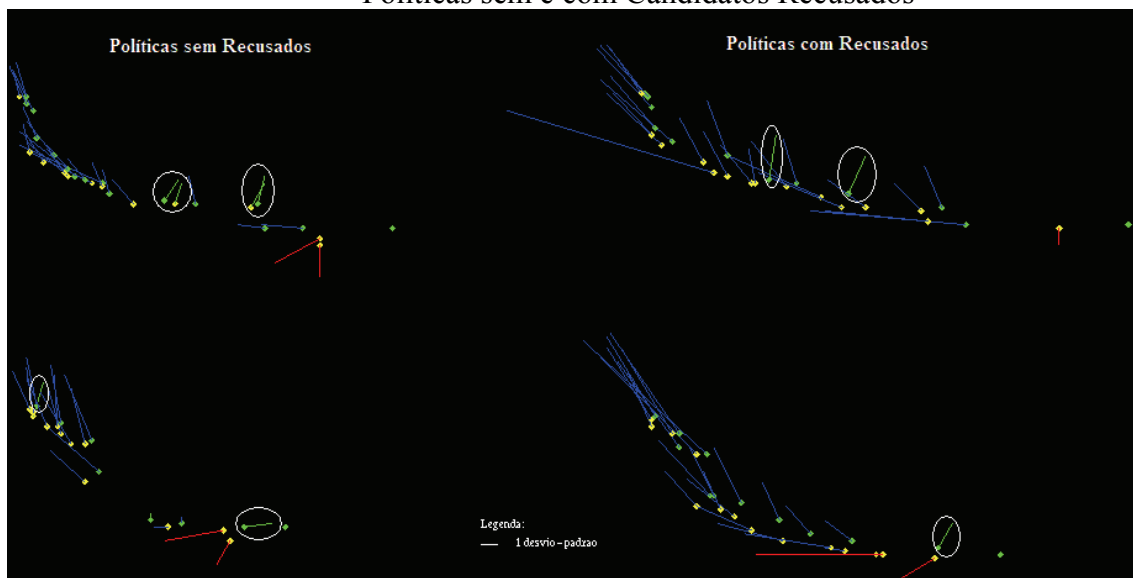


Fonte: elaborado pelos autores com dados do Banco do Povo

Buscamos uma medida que minimizasse a probabilidade de erro humano e criamos os denominados Vetores de Significância, já previamente definidos na seção metodológica. A FIG.5b apresenta o resultado dos vetores de significância para as possíveis políticas de crédito situadas nas duas camadas de Pareto (1º. e 2º. Níveis), pois, são aquelas que compõem a Fronteira de Pareto que merecem ser avaliadas. A cada política, em cada período  $t_0$  (abaixo) e  $t_1$  e (acima), e em cada grupo com (direita) ou sem (esquerda) recusados, associamos seu par (eficácia, eficiência) e seu vetor de significância  $vs_{x_0} = (dp_{x_0,ec_{x_0}})\hat{i} + (dp_{x_0,ef_{x_0}})\hat{j}$ , constituindo assim quatro campos vetoriais.

Representados graficamente para ambas as políticas (sem e com candidatos recusados), nos dois períodos ( $t_0$  e  $t_1$ ), os vetores verdes estão circulados na FIG.7b, e dentre eles, os relativos às políticas com candidatos recusados, no período  $t_1$ , são os mais significativos.

Figura 5b – Camadas de Paretos e seus Vetores de Significância para Políticas sem e com Candidatos Recusados



Fonte: elaborado pelos autores com dados do Banco do Povo

No gráfico esquerdo da FIG.6 temos o resultado da política selecionada em relação às políticas permutadas para fins de teste da hipótese nula. Esta política é a que apresenta o vetor de significância na cor verde situado mais à esquerda do gráfico direito superior da FIG.5b, e cujo desvio-padrão é ligeiramente superior ao do vetor verde situado mais à direita. Este segundo vetor de significância factível a ambos os objetivos pouco difere do primeiro. Ambas as políticas são definidas na TAB.2. A única distinção é que na primeira figura, cujo vetor de significância apresenta um desvio-padrão um pouco superior, só se seleciona as mulheres, enquanto na segunda seleciona-se também os homens. Outro aspecto interessante é o grau de escolaridade: em ambas, os clientes devem ter somente até o 1º grau de escolaridade completo.

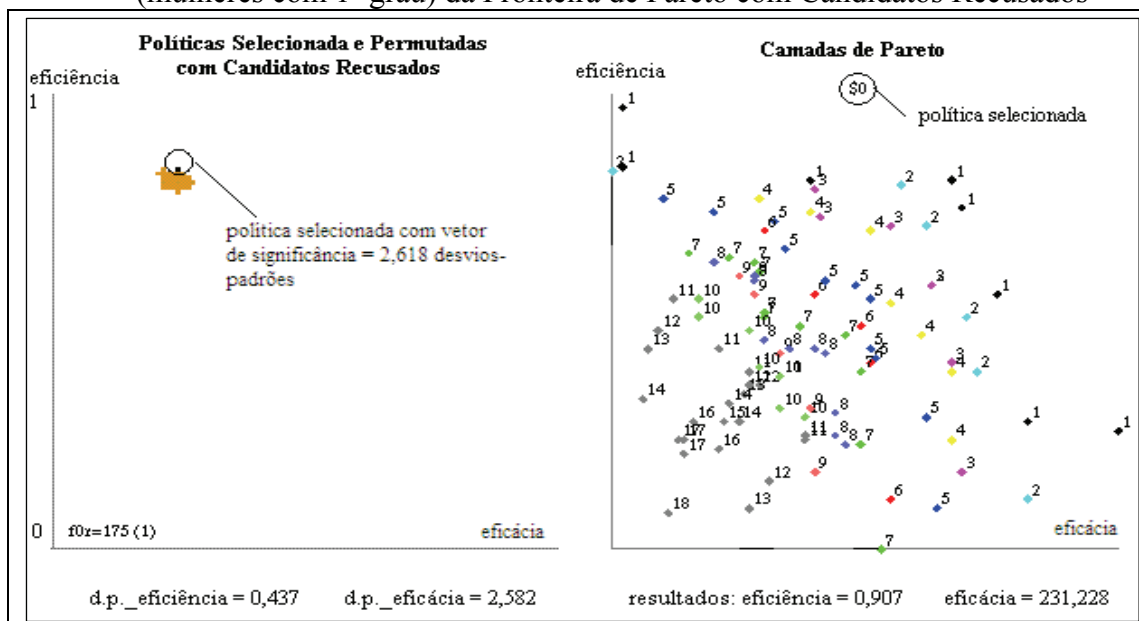
Podemos notar que seu posicionamento é claramente superior às políticas permutadas, como o gráfico à direita pode visualmente esclarecer. Neste último gráfico temos os pontos de todas as políticas (selecionada e as demais permutadas), e a política original é representada por um \$ antes do número da camada de Pareto a que pertence. Notamos que, de fato, trata-se de um ponto de dominância. A FIG.6

Tabela 2 – Políticas Mais Significativas da Fronteira de Pareto  
(inclui candidatos recusados, período  $t_1$ )

Variável	Vetor Significância Verde = 2,618 d.p. eficiência: 2,582 d.p. eficácia: 0,437 d.p.	Vetor Significância Verde = 2,490 d.p. eficiência: 2,335 d.p. eficácia: 0,865 d.p.
PAB	Belo Horizonte	Belo Horizonte
Estado civil	solteiros, casados e outros	solteiros e casados
Gênero	mulheres	homens e mulheres
Grau Escolaridade	1º grau completo ou incompleto	1º grau completo ou incompleto
Renda líquida	R\$0,00 e mais	R\$0,00 e mais
Finalidade Crédito	capital misto, fixo e giro	capital misto, fixo e giro
Valor do Crédito	todos os valores/faixas	todos os valores/faixas

Fonte: elaborado pelos autores

Figura 6 – Resultado do Melhor Vetor de Significância Política Seleccionada  
(mulheres com 1º grau) da Fronteira de Pareto com Candidatos Recusados



Fonte: elaborado pelos autores

É importante frisar que, em termos práticos, são os resultados que incluem os candidatos recusados os que merecem maior atenção e que, de fato, fornecem a resposta ao problema multi-objetivo aqui analisado. A razão para isto reside em dois aspectos: primeiro, a técnica escolhida da permutação cíclica, fundamental na solução do problema multi-objetivo deste trabalho, só pode ser aplicada corretamente se incluir o universo de candidatos a crédito da IMF. Caso contrário, a aleatoriedade gerada estaria “viesada” pela base dos clientes da IMF; e, segundo, que os candidatos recusados, que fazem parte do universo da IMF, mas que não estão incluídos na análise dos clientes, também não estão contemplados nos resultados das políticas exclusivas de clientes, não refletindo, assim, seu impacto sob os dois objetivos avaliados.

## 6. Conclusões e Trabalhos Futuros

Este trabalho apresentou uma técnica inovadora na área econômico-financeira para construir uma fronteira eficiente de política de seleção de crédito para instituições de microfinanças, baseada em otimização multi-objetivo para maximizar eficiência e eficácia. A principal novidade metodológica foi a aplicação da técnica de Monte Carlo de testes de permutação para gerar configurações sob a hipótese nula em problemas multi-objetivos, em substituição às técnicas econométricas clássicas, e às análises do tipo DEA e técnicas fuzzy,

mais frequentemente encontradas em artigos computacionais na área econômica, financeira e de engenharia. Para isso foi utilizado o novo conceito do vetor de significância, que permitiu realizar o cálculo de inferência da robustez dos resultados encontrados na fronteira de Pareto. A capacidade de selecionar a política de crédito mais adequada dentre aquelas contidas na fronteira ótima de Pareto, através de um critério quantitativo, sem que seja necessária a utilização de pesos arbitrariamente fixados para dosar a importância relativa das funções mono-objetivo de eficiência e eficácia, torna o método bastante atraente no contexto de avaliação de políticas de crédito em IMF.

Foram implementadas ferramentas de visualização que facilitam a interpretação dos resultados de forma interativa com o usuário, permitindo que se refinem as análises de políticas de crédito à medida que se compreende melhor o comportamento das políticas específicas analisadas. A utilização de uma técnica que fosse a mais próxima possível da distribuição empírica dos dados, minimizando a introdução de modelos estatísticos artificiais para os quais fosse difícil justificar sua aplicabilidade, permite que o próprio usuário possa conferir e criticar mais facilmente a robustez dos resultados obtidos.

A técnica aqui utilizada também pode se aplicar a políticas creditícias de bancos comerciais e, até mesmo, na avaliação de outras políticas, como, a tributária (buscando uma solução para o *trade-off* entre (i) maior nível de arrecadação e (ii) menores alíquotas tributárias, para que se possa definir qual a melhor estrutura tributária ser adotada), e as políticas públicas de cunho social (projetos de bolsa-escola, dentre outros), que também podem ser implementados buscando-se solucionar o *trade-off* entre custo mínimo e abrangência máxima.

## 7. Referências bibliográficas

- Chang, T.J., Meade, N., Beasley, J.E. Heuristics for Cardinality Constrained Portfolio Optimization. **Computers and Operations Research**, 27(13):1271-1302, 2000.
- Chankong, V., Haimes, Y.Y. **Multiobjective Decision Making: Theory and Methodology**. New York, USA. Dover Publications, 1983, 405p.
- Deb, K. **Multi-Objective Optimization using Evolutionary Algorithms**. London, John Wiley & Sons Ltd., 2001. 515p.
- Duczmal, L.H., Cançado, A.L.F., Takahashi, R.H.C. Delineation of Irregularly Shaped Disease Clusters Through Multiobjective Optimization. **Journal of Computational and Graphical Statistics**, v. 17, p. 243-262, 2008.
- Duczmal, L.H., Bessegato, L.F., Santos, M.A.C., Ferreira Neto, S.J. Introdução às Técnicas de Simulação em Estatística. **Relatório Técnico RTE-04/2003 Série Ensino**. Departamento de Estatística, UFMG, 2003.
- Ehrgott, M., Klamroth, K., Schwehm, C. An MCDM approach to portfolio optimization. **European Journal Of Operational Research**, 155(3):752-770, 2004.
- LARSON, H.J. **Introduction to Probability Theory and Statistical Inference**. John Wiley & Sons, 3ª ed., 1982. 637 pp.
- Manly, B.F.J. **Randomization, Bootstrap and Monte Carlo Methods in Biology**. Chapman & Hall, London, UK, 1997.
- Markowitz, H. Portfolio selection: efficient diversification of investments. **Journal of Finance**, 7(1):77-91, 1959.
- Mukerjee, A., Biswas, R., Deb, K., Mathur, A.P. Multi-objective evolutionary algorithms for the risk-return trade-off in bank-load management. **International Transactions in Operational Research**, 9(5):583-597, 2002.
- Mullei, S., Beling, P. Hybrid Evolutionary Algorithms for a Multiobjective Financial Problem. In: **Proceedings of the 1998 IEEE International Conference on Systems, Man, and Cybernetics**, 3925-3930. IEEE, October, 1998.

- Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P. **Numerical Recipes in C**. Cambridge University Press, USA, 2002. 994p.
- ROSS, S.M. **Simulation**. Academic Press, USA, 2002. 274pp.
- Ruspini, E.H., Zwir, I.S. Automated Qualitative Descriptions of Measurements. In: **Proceedings of the 16th IEEE Instrumentation and Measurement Technology Conference**, vol.(2), 1086-1091, Italy, 1999.
- Shoaf, J.S., Foster, J.A. A Genetic Algorithm Solution to the Efficient Set Problem: A Technique for Portfolio Selection Based on the Markowitz Model. In: **Proceedings of the Decision Sciences Institute Annual Meeting**, 571-573p., Orlando, 1996.

---

<sup>i</sup>Nas IMF, ao contrário dos bancos tradicionais, a variável taxa de juros é tratada *exogenamente* às demais variáveis avaliativas de seleção do crédito. Sua determinação é uma função direta do valor concedido de crédito: contratos de concessão de crédito abaixo de R\$960,00 (novecentos e sessenta reais) sofrem 2% a.m. de taxa de juros, enquanto valores superiores a R\$960,00 (novecentos e sessenta reais) sofrem correção de 3,9% a.m., e não há nenhum outro fator, especificamente atrelado ao risco de crédito, que possa interferir nestas taxas. Por esta razão, a taxa de juros não faz parte do modelo.

<sup>ii</sup> Entende-se por política creditícia todas as possibilidades de combinação entre as variáveis definidas como seletivas na determinação do crédito, ou seja, candidatos que geram os melhores resultados, simultaneamente, em termos de eficiência e eficácia.

<sup>iii</sup> Estas variáveis são as restrições definidas no problema multi-objetivo da seção 2.

<sup>iv</sup> Originalmente existia mais de 1 PAB e de categoria de renda líquida, na base de dados, e conseqüentemente a matriz P não tinha somente uma entrada na primeira e quinta variáveis. Com a finalidade de se mantermos a matriz P compatível com os dados originais e efetuarmos futuras modificações, optamos pela manutenção de todas as 7 (sete) variáveis.

<sup>v</sup> A distribuição de Poisson é usada com frequência para caracterizar o número de eventos que ocorre por unidade de tempo [Duczmal et al (2003: pp.28)].

<sup>vi</sup> A escolha de 10 clientes para cada candidato foi um critério que buscou minimizar o risco de se extrair da amostra indivíduos mais distantes e que, conseqüentemente, poderiam gerar uma maior variância no histórico de crédito, não assegurando, assim, a proximidade da amostra. Enfim, buscou-se manter o compromisso entre tamanho e proximidade amostral.

<sup>vii</sup> Para o presente trabalho, escolhemos o gerador ran1, o qual é capaz de remover correlações seriais de baixa ordem (solicitação). Sobre geradores aleatórios, ver Press et al (2002).