

**PLANEJAMENTO DE UMA FAZENDA EM
CONDIÇÕES DE RISCO: PROGRAMAÇÃO
LINEAR E SIMULAÇÃO
MULTIDIMENSIONAL**

CARLOS TADEU DOS SANTOS DIAS

Engenheiro Agrônomo

Orientador: Prof. Dr. CASSIO ROBERTO DE MELO GODOI

Tese apresentada à Escola Superior de
Agricultura "Luiz de Queiroz", da Uni-
versidade de São Paulo, para obtenção
do título de Doutor em Agronomia, Área
de Concentração: Estatística e Experi-
mentação Agronômica.

P I R A C I C A B A
Estado de São Paulo - Brasil
Janeiro - 1996

Dados Internacionais de Catalogação na Publicação (CIP)
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - Campus "Luiz de Queiroz"/USP

Dias, Carlos Tadeu dos Santos

Planejamento de uma fazenda em condições de risco: programação linear e simulação multidimensional / Carlos Tadeu dos Santos Dias. - - Piracicaba, 1996. 100p. : il.

Tese (doutorado) - - Escola Superior de Agricultura Luiz de Queiroz, 1996. Bibliografia.

1. Empresa agrícola - Risco - Planejamento 2. Programação linear 3. Simulação
I. Título

CDD 338.1
519.72
511.8

Estatística, para a maioria das pessoas, lembra grandes quantidades de números, gráficos complicados, etc. Entretanto como define “Wallis e Robert” a estatística é um conjunto de métodos para se tomar decisões satisfatórias na presença da incerteza.

Aos meus pais

Ubaldo Rocha Dias e Francisca dos Santos Dias

Dedico.

*Aos meus irmãos,
Francisco, Ana, Silene e Verônica
Ofereço.*

AGRADECIMENTOS

Ao Professor Cassio Roberto de Melo Godoi, pela orientação, apoio e amizade dedicada.

Ao Professor Fernando Curi Peres, por nos ter colocado esse desafio.

Ao Professor Rodolfo Roffmann pelas sugestões.

Aos Professores do IME/USP, em especial aos Professores Dalton e Galvão, pela contribuição na realização deste trabalho.

À Celia, pelo companheirismo e valiosa ajuda nesta árdua tarefa.

Aos Professores do Departamento de Matemática e Estatística da ESALQ/USP, pelos ensinamentos e convívio.

Aos Colegas de curso, César, Gener, Joel, Rosana e Rui, pelo companheirismo e amizade.

Aos funcionários do Departamento de Matemática e Estatística da ESALQ/USP, Expedita, Jorge, Luciane, Rosa, Rosni e Solange.

Aos Professores do Departamento de Bioestatística da UNESP/Botucatu, pela confiança e oportunidade que nos deram para realização deste curso.

Aos Professores e funcionários do Departamento de Horticultura, em especial ao Professor Keigo Minami.

Aos Professores Valdeci Bissera e José Henrique, da Universidade Federal do Ceará, pelo fornecimento dos dados.

À Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior (CAPES), através do seu Programa Institucional de Capacitação de docentes (PICD), pelo apoio financeiro.

A todos aqueles que, direta ou indiretamente, colaboraram com esta importante etapa da minha vida.

SUMÁRIO

| | | |
|----------|--|-----------|
| 1 | INTRODUÇÃO | 1 |
| 1.1 | Importância do Problema | 1 |
| 1.2 | Pressuposições e objetivo | 3 |
| 2 | A PROGRAMAÇÃO LINEAR COMO INSTRUMENTO DE PLANEJAMENTO DA EMPRESA AGRÍCOLA | 4 |
| 2.1 | Origens e Usos | 4 |
| 2.2 | Principais Conceitos na Representação de um Problema de Programação Linear | 5 |
| 2.2.1 | Aspectos Gerais da Programação Linear | 5 |
| 2.3 | Terminologia para Soluções do Modelo | 7 |
| 2.4 | Suposições da Programação Linear | 8 |
| 3 | A ANÁLISE DA DECISÃO | 14 |
| 3.1 | Histórico | 14 |
| 3.2 | Natureza da Análise da Decisão | 16 |
| 3.3 | A Análise da Decisão na Agricultura | 18 |
| 3.4 | Abordagem Normativa e Descritiva da Decisão | 19 |
| 3.5 | Modelos de Decisão | 21 |
| 3.5.1 | Critério de decisão em caso de certeza | 24 |
| 3.5.2 | Critérios de decisão sob estrita incerteza | 24 |
| 3.5.3 | Critérios de decisão sob risco | 28 |
| 4 | MODELOS COM INCLUSÃO DE INCERTEZA OU RISCO EM PROBLEMAS DE PROGRAMAÇÃO LI- | |

| | |
|--|-----------|
| NEAR | 30 |
| 4.1 Modelo com Base no Critério Minimax de Savage | 33 |
| 4.2 Maximização da Margem Bruta Total com Restrições de Máxima Perda Admissível | 34 |
| 4.3 Inclusão de Risco através da Programação Estocástica . . . | 36 |
| 4.4 Inclusão de Risco através da Aplicação Simultânea de Téc- nicas de Simulação e Programação Linear | 38 |
| 5 GERAÇÃO DE VARIÁVEIS ALEATÓRIAS MULTI- VARIADAS | 43 |
| 5.1 Métodos de Geração | 43 |
| 5.2 Distribuições Bivariadas Correlacionadas | 47 |
| 6 METODOLOGIA PROPOSTA | 56 |
| 6.1 Adoção de Funções Probabilísticas para os Coeficientes A- leatórios | 56 |
| 6.2 Uso Simultâneo de Métodos de Simulação e Programação Linear. | 57 |
| 6.3 Métodos de Obtenção de Valores Amostrais para os Coe- ficientes Aleatórios do Problema de Programação Linear. . | 58 |
| 6.4 Obtenção e Análise dos Resultados | 68 |
| 6.5 Algoritmo Computacional do Método Proposto | 69 |
| 7 APRESENTAÇÃO DO CASO ESTUDADO | 71 |
| 8 CONCLUSÕES E SUGESTÕES | 85 |
| REFERÊNCIAS BIBLIOGRÁFICAS | 87 |

PLANEJAMENTO DE UMA FAZENDA EM CONDIÇÕES DE RISCO: PROGRAMAÇÃO LINEAR E SIMULAÇÃO MULTIDIMENSIONAL

Autor: CARLOS TADEU DOS SANTOS DIAS

Orientador: Prof. Dr. CASSIO ROBERTO DE MELO GODOI

RESUMO

Este estudo teve como objetivo propor um método de inclusão de risco em modelos de planejamento da empresa agrícola, através da utilização de técnicas de simulação e programação linear, conjuntamente, levando-se em consideração as correlações entre os coeficientes do modelo de programação linear.

Foram utilizados dados de um modelo de exploração do Perímetro de Irrigação Senador Nilo Coelho (BA/PE), como aplicação do método proposto, para estimar os parâmetros da distribuição da margem bruta e os planos propostos pelo modelo de programação linear.

Os resultados encontrados mostram um risco econômico maior, traduzido em termos de variância da margem bruta, quando se utiliza o modelo com correlações entre os coeficientes. Os planos modais propostos pelos modelos de programação linear, utilizando simulação com e sem dependência não diferem muito do plano proposto pelo modelo determinístico (sem simulação), mas diferem do plano atual.

FARM PLANNING UNDER RISK: LINEAR PROGRAMMING AND MULTIVARIATE SIMULATION

Author: CARLOS TADEU DOS SANTOS DIAS

Adviser: Prof. Dr. CASSIO ROBERTO DE MELO GODOI

SUMMARY

This study had the objective of proposing a method of inclusion of risk factors in a farm planning through simulation techniques and linear programming combined, taking in consideration the correlation coefficients of the sampled coefficient of the linear programming model.

Data from Perímetro de Irrigação Senador Nilo Coelho (BA/PE) were used to estimate the gross margin distribution parameters and the proposed plans for the linear programming model.

The results pointed out to a bigger economic risk as in the variance of gross margin when the model of correlation coefficients was applied. The proposed plans for the linear programming model, with simulations that included and did not included dependency, did not differ from the proposed plan, but they did differ from the current plan.

1 INTRODUÇÃO

1.1 Importância do Problema

O crescimento econômico brasileiro passa pela produção adequada de alimentos, fibras e outros produtos agrícolas. O volume de produtos agrícolas é função da decisão e da racionalidade do agricultor, que depende, por sua vez, de métodos eficientes que o auxiliem na otimização da sua decisão.

O agricultor, ao procurar otimizar a sua decisão, deve escolher, dentre as alternativas de produção disponíveis, a mais eficiente na utilização de recursos produtivos e a que satisfaça a certos objetivos preestabelecidos. Numa situação estática e com conhecimento perfeito, esses objetivos podem se resumir à maximização do lucro ou minimização dos custos. Mesmo com conhecimento perfeito, dificilmente a decisão dos agricultores, especialmente dos pequenos, seria a de simplesmente, maximizar lucros esperados. Sckstowisky ¹, citado por RODRIGUES (1987), demonstrou que, quando o mesmo indivíduo toma as decisões sobre os investimentos e consumo, então, só por acaso, acontecerá a coincidência entre a maximização da utilidade com a maximização do lucro.

A programação matemática é um instrumento fundamental do planejamento agrícola, auxiliando o agricultor a formular os problemas e obter a melhor solução.

A programação linear, um dos métodos da programação matemática, tem encontrado maior aplicação. Entretanto, a formulação original dos problemas de programação linear implica na adoção de valores determinados para coefi-

¹SCKSTOWISKY, T. Welfare & competition. Homewood, Richard D. Irvin, Inc., 1951. 142-

cientes (constantes do modelo de programação linear) que, na realidade, são variáveis aleatórias. A sua utilização, na forma determinista, no planejamento agrícola fica, portanto, limitada se for considerado que a empresa agrícola atua sob condições de risco econômico e que a aleatoriedade de determinados coeficientes, responsáveis por essa situação, não pode ser desprezada.

Alguns dos métodos utilizados para expressar o comportamento aleatório dos coeficientes, devido à sua especificidade, exigem que os coeficientes aleatórios estejam localizados na função objetivo do modelo de programação linear, e que se suponha normal a distribuição de probabilidades da variável a ser otimizada. Outros métodos apresentam a desvantagem de não permitir a existência de uma eventual correlação entre os coeficientes aleatórios. Outros ainda definem uma função utilidade do agricultor restrita a poucas situações reais, ou não permitem a inclusão de coeficientes com diferentes distribuições de probabilidades, ou se restringem a situações muito específicas de aversão ao risco.

O conceito de utilidade não é baseado em algum tipo de satisfação psicológica por parte do agricultor, mas no seu comportamento de escolha. A hipótese de que um agricultor obtém mais satisfação ou utilidade com A do que com B significa que ele preferiria A, em vez de B, se A e B fossem apresentados como alternativas. Além disso, supõe-se que o agricultor faça escolhas consistentes, de forma que se A é preferido a B e B é preferido a C então A é preferido a C. A moderna teoria da utilidade supõe apenas uma escala de utilidade ordinal; isto é, para todos os pares de alternativas A e B, supõe-se que o agricultor tenha preferência por A ou por B, ou que ele seja indiferente a escolha. Não se supõe que ele seja capaz de atribuir números que representem o grau de utilidade de cada artigo. Uma classificação (ordem de preferência) de artigos do agricultor é expressa matematicamente pela sua função de utilidade. Esta associa certos números (utilidade) com várias quantidades de bens consumidos ou produzidos, mas estes números representam apenas uma classificação ou ordenação de preferências. Portanto, se utilidade da alternativa A é 20 e a utilidade da alternativa B é 5, então A é preferido a B, mas é incorreto dizer

que A seja preferido 4 vezes mais do que B (WEBER, 1977).

Apesar da existência de diversos métodos de inclusão de risco em modelos de programação linear voltados para o planejamento da empresa agrícola, pouco se tem conseguido em termos de aplicabilidade a situações onde possa ser feita qualquer pressuposição acerca da distribuição conjunta de probabilidades dos coeficientes considerados aleatórios.

1.2 Pressuposições e objetivo

- A instabilidade de determinados fatores incontrolláveis, como clima (chuva, seca), interferências governamentais, preços e ocorrência de pragas e doenças, é o principal responsável pela situação de risco sob a qual atua a empresa agrícola;
- O risco envolvido na atividade rural influencia o comportamento do agricultor;
- Os modelos de programação linear, ao incluírem, adequadamente, determinados coeficientes aleatórios, refletirão a situação de risco enfrentada pelo agricultor.
- Não se dispondo de dados históricos suficientes, a distribuição de probabilidades dos coeficientes aleatórios pode ser definida a partir de probabilidades estimadas de maneira subjetiva e fornecidas pelo agricultor.

Objetivou-se, com este trabalho, propor um método de inclusão de risco em modelos de planejamento da empresa agrícola, através da utilização de técnicas de simulação e programação linear, conjuntamente, levando-se em consideração relações de dependência entre os coeficientes. Avaliou-se as vantagens apresentadas por esse modelo.

2 A PROGRAMAÇÃO LINEAR COMO INSTRUMENTO DE PLANEJAMENTO DA EMPRESA AGRÍCOLA

2.1 Origens e Usos

A programação linear é um método matemático, frequentemente útil em situações de tomada de decisão, onde se requer a escolha de uma única alternativa dentre um conjunto de possibilidades. Foi utilizado, inicialmente, durante a 2^a. Grande Guerra, como método de obtenção de rotas que minimizassem os custos de transportes dos Aliados, na distribuição de tropas, suprimento de armas, munições, alimentos, medicamentos e demais itens que compunham os elementos logísticos militares.

Com o fim da guerra, o método tornou-se mais refinado e tem sido usado como instrumento de pesquisa também, por economistas agrícolas na otimização e organização dos recursos e empreendimentos rurais, sugerindo ajustes desejáveis, visando maximização dos lucros na produção de produtos agropecuários, pela minimização de custos no processamento de produtos como fertilizantes e rações, obtendo modelos de distribuição espacial equilibrados quanto ao fluxo de produtos agrícolas, indicando volumes ótimos de fluxos inter-regionais de fatores de produção e produtos específicos.

De fato, a sua aplicação se estende a qualquer tipo de problema que permita a definição de um objetivo em termos quantitativos. A quantidade a ser maximizada poderia, por exemplo, representar o lucro de uma fazenda, unidades de alimentos de uma determinada área agrícola ou calorias e outros componentes

alimentares produzidos por um segmento do setor agrícola. A quantidade a ser minimizada poderia, por exemplo, ser o custo de uma ração, de transporte entre origens e destinos de um produto, de armazenamento da produção de grãos, ou da quantidade de trabalho necessária para providenciar as necessidades alimentares de uma determinada população.

Sumariamente, a programação linear trata tipicamente com o problema de alocação de recursos limitados a atividades em competição, da melhor maneira possível, isto é, ótima (HILLIER & LIEBERMAN 1988).

Este problema de alocação pode aparecer toda vez que alguém precise selecionar o nível de certas atividades que competem por recursos escassos necessários para desempenhá-las.

2.2 Principais Conceitos na Representação de um Problema de Programação Linear

2.2.1 Aspectos Gerais da Programação Linear

A programação linear usa um modelo matemático para descrever o problema em questão. O adjetivo “linear” significa que é requerido que todas as funções matemáticas neste modelo, sejam funções lineares. A palavra “programação” aqui não se refere à programação de computadores; ao contrário, trata-se essencialmente de um sinônimo de planejamento.

O termo linear não é tão restritivo quanto pode parecer. Significa que os coeficientes usados expressam um comportamento constante ou representam uma relação linear entre, por exemplo, o fator de produção e o produto final e os preços pagos pelos recursos ou recebidos pelos produtos são constantes. Entretanto, isto não quer dizer que retornos decrescentes não possam ser incorporados aos problemas de programação linear.

A programação linear utiliza procedimentos matemáticos baseados em relações lineares e inequações. As inequações irão permitir que não se exija necessariamente o uso integral dos recursos disponíveis, e que a quantidade de qualquer atividade explorada ou produto seja igual ou maior que zero. Desta forma, assegura-se que a quantidade usada de qualquer recurso seja menor ou igual à quantidade disponível e que a produção seja maior ou igual a zero.

O modelo de programação linear com n variáveis ou atividades e m restrições pode ser formulado da seguinte forma:

Maximizar ou Minimizar

$$Z = \sum_{j=1}^n c_j x_j \quad (1)$$

Sujeito a:

$$\sum_{j=1}^n a_{ij} x_j \leq = \geq b_i, \quad (i = 1, 2, \dots, m) \quad (2)$$

$$x_j \geq 0, \quad (j = 1, 2, \dots, n) \quad (3)$$

onde,

Z - função objetivo que será maximizada ou minimizada. É considerada a medida geral de eficácia;

x_j - nível da j -ésima atividade (uma variável de decisão);

c_j - constante, representando a contribuição ou custo por unidade da j -ésima atividade. Genericamente, c_j é o aumento em Z que resultará de cada unidade acrescentada em x_j ;

a_{ij} - quantidade constante do recurso i consumida em cada unidade de atividade j ;

b_i - constante, representando a limitação ou disponibilidade do i -ésimo recurso.

A notação $\leq = \geq$ significa que as restrições são inequações do tipo (\leq), (\geq) ou igualdades ($=$), e nesse último caso, impõe-se que todo o recurso disponível seja totalmente utilizado. Essas são as restrições funcionais, enquanto que as restrições $x_j \geq 0$ são chamadas restrições de não negatividade.

As constantes de entrada, a_{ij} , b_i e c_j podem ser chamadas de parâmetros do modelo. Nos problemas de planejamento da empresa rural, a função objetivo deve, necessariamente, ser expressa em termos físicos ou monetários, dependendo do tipo de problema analisado.

Estabelecida a função objetivo, a programação linear terá como alternativas escolher métodos ou processos disponíveis para atingi-la. Tais alternativas apresentam-se, em geral, como técnicas disponíveis de produção, culturas aptas para exploração na propriedade, ou fontes de produção capazes de contribuir para o objetivo da análise.

Obviamente, um problema de programação linear não existiria se os recursos não fossem limitados. Para a maioria dos problemas existem restrições que estabelecem limites para os planos em consideração. Uma propriedade rural sempre apresenta uma quantia limitada de terra, de construções, de máquinas, de trabalho e de capital. Enquanto os recursos fixos representam o principal tipo de restrição, outras restrições subjetivas ou institucionais também podem existir.

A programação linear somente será um meio eficiente de se determinar planos ótimos se existirem diversas alternativas e diversas restrições. De outra forma, a maximização do lucro ou minimização dos custos se tornaria trivial, tornando mais recomendável a utilização de simples métodos orçamentários de cálculo.

2.3 Terminologia para Soluções do Modelo

O termo solução é habitualmente usado como a resposta final a um problema, porém, em programação linear, é bem diferente. Aqui, qualquer especificação de valores para as variáveis de decisão x_j , é chamada de solução, não importando se se trata de uma escolha desejável ou mesmo permissível.

Identifica-se, portanto, diferentes tipos de soluções, usando-se um adjetivo apropriado, como segue.

Solução viável é uma solução em que todas as restrições são satisfeitas. Quando, pelo menos uma restrição não é satisfeita, a solução é dita inviável. É possível que um problema não tenha nenhuma solução viável, e outro tenha infinitas soluções viáveis.

Se há soluções viáveis, o intento da programação linear é descobrir qual é a “melhor”, de acordo com o valor da função objetivo.

Uma solução ótima é uma solução viável que tem o valor mais favorável da função objetivo. Por valor mais favorável, entende-se o maior ou menor valor, dependendo do objetivo ser maximizar ou minimizar. Portanto, uma solução ótima maximiza ou minimiza a função objetivo por toda a região viável (pontos que se encontram dentro ou nos limites da região delimitada pelas restrições).

Frequentemente, um problema terá apenas uma solução ótima. Entretanto, é possível haver múltiplas soluções ótimas.

Se as variáveis forem contínuas, qualquer problema com múltiplas soluções ótimas terá um número infinito delas

Outra possibilidade é que o problema não tenha soluções ótimas. Isto ocorre se não houver soluções viáveis, ou as restrições não impedirem o crescimento indefinido do valor da função objetivo na direção (positiva ou negativa) favorável.

2.4 Suposições da Programação Linear

As suposições da programação linear estão implícitas na formulação do modelo 1,2,3 da seção 2.2. No entanto, deve-se ressaltar estas suposições para que se possa avaliar mais facilmente quão bem a programação linear se aplica a qualquer problema dado. Se essas suposições não forem satisfeitas ou não se aplicarem ao problema estudado, a programação linear não proporcionará uma representação satisfatória do problema.

Certeza: A suposição de certeza é que todos os parâmetros do modelo, os valores a_{ij} , b_i , e c_j , são constantes conhecidas. Em problemas reais, como o do planejamento rural, esta suposição raramente é precisamente satisfeita. Os modelos de programação linear usualmente são formulados no sentido de selecionar algum curso de ação futura. Por isso os parâmetros usados são baseados numa predição de condições futuras, as quais introduzem inevitavelmente algum grau de incerteza.

Nesse caso, é importante realizar uma análise de sensibilidade completa após ter encontrado a solução de programação linear com os valores supostos dos parâmetros. O propósito geral é identificar os parâmetros relativamente sensíveis, isto é, aqueles cuja variação causa modificação significativa da solução ótima e, então, selecionar a solução que permaneça boa por toda a extensão de valores prováveis dos parâmetros sensíveis.

Proporcionalidade: é uma suposição sobre atividades individuais consideradas independentemente uma das outras (enquanto que a suposição subsequente de aditividade se refere ao efeito de conduzir atividades conjuntamente). Assim, considere o caso em que apenas uma das n atividades é tratada. Chame-a de atividade k , de modo que $x_j = 0$ para todo $j = 1, 2, \dots, n$, exceto $j = k$.

A suposição é que, neste caso, a medida de eficácia Z é igual a $c_k x_k$ e o uso de cada recurso i é igual a $a_{ik} x_k$, isto é, ambas as quantidades são diretamente proporcionais ao nível da atividade k . Isto implica, especificamente, que não existe nenhuma carga extra de inicialização da atividade e que a proporcionalidade se mantém por toda a extensão dos níveis da atividade.

O caso de atividades com valores positivos e negativos é um tipo importante de não proporcionalidade, mas que assim mesmo pode ser tratado pela programação linear, reformulando-se o problema apropriadamente (HILLIER & LIEBERMAN, 1988).

Segundo esses autores, às vezes é necessário lidar com atividades às quais é permitido serem positivas ou negativas. Quando não existe limite para os

valores negativos permitidos, cada uma destas atividades (x_j) pode ser substituída, em todo o modelo, pela diferença de duas novas variáveis não-negativas (x_j^+ e x_j^-), de tal modo que

$$x_j = (x_j^+ - x_j^-),$$

onde, $x_j^+ \geq 0$, representa os valores positivos de x_j e $x_j^- \geq 0$, os seus valores negativos. Especificamente

$$x_j^+ = \begin{cases} +x_j, & \text{se } x_j \geq 0 \\ 0, & \text{se } x_j \leq 0, \end{cases}$$

$$x_j^- = \begin{cases} 0, & \text{se } x_j \geq 0 \\ -x_j, & \text{se } x_j \leq 0, \end{cases}$$

para todas as soluções básicas viáveis, uma vez que tais soluções têm, necessariamente, a propriedade de que $x_j^+ = 0$ ou $x_j^- = 0$ (ou ambos). O efeito da escolha de valor para x_j pode ser bastante diferente para valores positivos e negativos. Por exemplo, suponha-se que x_j represente o nível de estoque de um produto particular. Se $x_j > 0$ (de modo que $x_j^+ > 0$ e $x_j^- = 0$), os custos que incorrerão vão incluir as despesas de estoque e encargos de juros sobre o capital aplicado neste estoque. Por outro lado, $x_j < 0$ (de modo que $x_j^- > 0$ e $x_j^+ = 0$) significa que ocorreu uma falta de x_j^- . Os custos, neste caso, resultam da perda de vendas, tanto agora (se os clientes não esperarem) quanto no futuro (se os clientes descontentes não voltarem). Por causa desta diferença entre os casos positivo e negativo, o custo de x_j não é simplesmente proporcional a x_j . Portanto a suposição de proporcionalidade da programação linear é violada para este exemplo se for utilizada diretamente a variável x_j .

Felizmente, enquanto a suposição de proporcionalidade se mantiver para os casos positivo e negativo considerados separadamente, a função objetivo poderá ser reformulada num formato de programação linear usando-se x_j^+ e x_j^- .

Seja Z_j , a contribuição de x_j para a função objetivo Z . Para constantes apropriadas, c_j^+ e c_j^- , se

$$Z_j = \begin{cases} c_j^+ x_j, & \text{para } x_j \geq 0 \\ c_j^- (-x_j), & \text{para } x_j \leq 0, \end{cases}$$

então,

$$Z_j = c_j^+ x_j^+ + c_j^- x_j^-,$$

Um caso particular e importante é quando $c_j^+ = c_j^-$ (valor comum de c_j), de modo que Z_j é simplesmente proporcional ao valor absoluto de x_j , $|x_j|$. Note-se que

$$|x_j| = x_j^+ + x_j^-$$

Por isso, se

$$Z_j = c_j |x_j|$$

então,

$$Z_j = c_j (x_j^+ + x_j^-).$$

Outra situação que exige métodos especiais é o problema do custo fixo de preparação. Em tais casos, o custo total da atividade é a soma de um custo variável relacionado ao nível da atividade e um custo de preparação requerido para iniciar a atividade. Frequentemente, o custo variável será, pelo menos, aproximadamente proporcional ao nível da atividade. Assim, se x_j denotar o nível da atividade j , o custo total aproximado da atividade j frequentemente será $(k_j + c_j x_j)$ se $x_j > 0$, enquanto que o custo será zero se $x_j = 0$. Se não fosse pelo custo de preparação k_j , esta estrutura de custo sugeriria a possibilidade de uma formulação de programação linear para determinar os níveis ótimos das atividades em competição. No entanto, mesmo com os k_j , a programação linear inteira ² ainda pode

²Os problemas de programação inteira surgem frequentemente porque algumas ou todas as variáveis de decisão têm que estar restritas a valores inteiros.

ser usada, reformulando-se o problema. A reformulação envolve a introdução de n decisões sim-ou-não quanto a se realizar ($x_j > 0$) as n respectivas atividades, e, então, representar cada uma destas decisões, por uma variável binária y_j . Assim,

$$Z = \sum_{j=1}^n (c_j x_j + k_j y_j)$$

onde

$$y_j = \begin{cases} 1, & \text{se } x_j > 0 \\ 0, & \text{se } x_j = 0 \end{cases}$$

Aditividade: A suposição de proporcionalidade não é suficiente para garantir que a função objetivo e as funções de restrição sejam lineares. “Os termos de produtos cruzados” irão ocorrer se houver interação entre algumas das atividades que mudem a medida total de eficácia ou o uso total de algum recurso. A aditividade supõe que não existe uma tal interação entre quaisquer das atividades. Por isso, a suposição de aditividade requer que, dados quaisquer níveis de atividade (x_1, x_2, \dots, x_n) , o uso total de cada recurso e a medida total de eficácia resultante sejam iguais à soma das quantidades correspondentes geradas por toda a atividade conduzida pelo recurso. Desta forma, admite-se produtos marginais físicos constantes, em desobediência à lei dos rendimentos físicos marginais decrescentes.

Divisibilidade: Há casos em que variáveis de decisão só tem significação física se tiverem valores inteiros. Entretanto, a solução obtida pela programação linear geralmente não é inteira. Por isso, a suposição de divisibilidade é que as unidades de atividade possam ser divididas em qualquer nível fracional, para que sejam permissíveis valores não-inteiros para as variáveis de decisão.

A programação linear ainda é frequentemente aplicada, mesmo quando é requerida uma solução inteira. Se a solução obtida não for inteira, as variáveis não-inteiras são simplesmente arredondadas para valores inteiros. Isto pode ser satisfatório (particularmente, se as variáveis de decisão forem grandes), mas pode haver certos perigos, e se essa abordagem não puder ser usada, a pro-

gramação linear inteira deve ser considerada. De qualquer forma, existem diversos fatores de produção agrícola que podem ser usados em quantias fracionárias.

Finitude: Admite-se que o número de atividades alternativas e restrições sejam limitadas. É uma pressuposição prática e realista, pois, em todos os casos o interesse se restringe a um número comparativamente pequeno de atividades em competição ou alternativas.

3 A ANÁLISE DA DECISÃO

3.1 Histórico

A teoria da decisão estuda metodologias e estruturas analíticas que servem de apoio para os administradores tomarem decisões cada vez mais complexas.

Os seres humanos tem tomado decisões há milhares de anos pelo uso do senso comum e métodos heurísticos, porém, a mais antiga contribuição para o desenvolvimento da análise da decisão na sua forma atual foi dada por Daniel Bernouilli em 1738, o qual contestou ser o dinheiro uma medida adequada de valor. Surgiu então o conceito de utilidade, medida numérica para descrever a real importância das consequências de uma decisão. Bernouilli mostrou que para o mesmo valor em dinheiro, a utilidade desse mesmo dinheiro variava de indivíduo para indivíduo, devendo-se considerar o seu valor moral, o que na terminologia moderna é essencialmente a sua utilidade esperada. Embora a idéia de utilidade fosse boa naquela época, a sua aplicação foi difícil, por não ser possível medir realmente e de maneira satisfatória a função-utilidade de uma pessoa.

Nos séculos XVIII, XIX e início do XX, a teoria da utilidade desenvolveu-se de modo confuso. Conceitos, tais como, funções diferenças de valor, funções de valores ordinais e funções-utilidades foram embaralhadas, confundidas. Diversos economistas da década de 1920 argumentaram que a teoria econômica poderia ser fundamentada com hipóteses menos fortes acerca da medida de utilidade. Ao invés de proporem hipóteses de que um número representava o valor de qualquer conjunto de bens (seria a chamada utilidade cardinal), podia-se simplesmente assumir que, dados dois conjuntos de bens, era possível dizer qual deles era o preferido (dando origem à chamada utilidade ordinal).

Esse assunto, de raízes antigas assumiu formas modernas devido a RAMSEY (1964), NEUMANN & MORGENTERN (1953) e SAVAGE (1954), tomando corpo principalmente após a Segunda Grande Guerra, época da Pesquisa Operacional e da Ciência da Administração. Aprofundaram-se as pesquisas a respeito de probabilidades subjetivas, análise de dominância estocástica, teoria da utilidade com multiatributos, e tomada de decisões coletivas sob a rubrica de teoria da escolha social. Mais ainda, desafios empíricos aos axiomas considerados básicos deram origem a novos modelos, resultando na axiomatização de outras teorias que tentavam acomodar os resultados experimentais sem perder muito da elegância matemática.

Na década de 60, importantes contribuições foram feitas ao aplicar-se a chamada Teoria Estatística da Decisão a Problemas de Economia. Criou-se o termo “Análise da Decisão”, mais abrangente para o assunto que se desenvolvia de forma considerável.

As últimas décadas têm sido promissoras para o desenvolvimento da teoria e da análise da decisão, e em 1988 comemorou-se 250 anos da publicação do artigo “Specimen Theoriae novae de mensura sortis” de Daniel Bernouilli, marco inicial da Teoria da Decisão. Embora experimentos continuem a demonstrar persistentes e sistemáticas violações das premissas geralmente aceitas da teoria da utilidade, novas teorias têm se preocupado em acomodar violações predizíveis dos modelos tradicionais, sem abandonar muito da elegância matemática e do poder analítico do conceito de utilidade, conservando implicações com respeito à atitude quanto a risco, dominância estocástica e outros fatores, e permitindo padrões de preferência que muitas pessoas consideram razoáveis.

3.2 Natureza da Análise da Decisão

Pode-se dizer que problema é uma situação difícil ou duvidosa, e decisão é o processo de resolver tal situação. Howard ³, citado por SILVA (1990), afirma que “tomar uma decisão é aquilo que você faz quando não sabe o que fazer”. A análise da decisão é, simplesmente, o termo que se refere ao estudo cuidadoso que precede a uma decisão.

As decisões podem ser classificadas em programadas, que são aquelas rotineiras, repetitivas e, normalmente diretas; e não programadas, que são aquelas decisões únicas e complexas, não podendo ser trabalhadas de uma maneira intuitiva. Os problemas de programação linear se encaixam nesse último tipo de decisão. Elas são consideradas as mais difíceis por terem uma ou mais das seguintes características: impacto ao longo do tempo, existência de inúmeras alternativas, influência em outras pessoas além do tomador de decisão, consequências incertas e dimensões múltiplas de valor, embora não sejam necessariamente características da programação linear.

Ao se pensar acerca dos processos de tomada de decisão, deve-se proceder segundo o senso comum de que, na maioria das vezes, bons processos de decisão levariam a melhores resultados que piores processos, e pode-se afirmar que uma análise racional e sistemática contribui para a tomada de boas decisões. Mais ainda, não se deseja substituir os processos intuitivos de resolução de problemas por processos puramente analíticos. Ao contrário, busca-se combinar modos analíticos e modos intuitivos de tal sorte que a análise formal dê suporte à intuição.

Novos campos de estudo estão se desenvolvendo rapidamente, como por exemplo, a Inteligência Artificial e os Sistemas de Suporte à Decisão, os quais combinam a habilidade criativa do ser humano com as habilidades lógicas e de processamento da informação dos computadores, permitindo a ligação entre os modelos de decisão e uma grande massa de dados. A análise da decisão segue o seguinte

³HOWARD, R.A. The foundations of decision analysis. **IEEE Transactions on Systems, Science and Cybernetics**, SSC-4, n.3, set. 1968.

fluxo: 1. Formular o problema; 2. Construir o modelo; 3. Resolver o modelo; 4. Interpretar o resultado e tirar conclusões a respeito do problema; 5. Fazer uma análise de sensibilidade (usando-a para reformular o problema ou alterar o modelo) e 6. Implementar a solução.

Costuma-se considerar a análise da decisão como uma série de técnicas estatísticas (a chamada Teoria Estatística da Decisão). Enquanto estas técnicas desempenham um papel importante na solução de problemas, a visão atual é que na prática, os decisores lucram mais da análise em si mesma do que das respostas que ela produz. O objetivo é identificar uma decisão que se espera ser aquela que melhor satisfaça os objetivos especificados por quem deseja o problema resolvido.

Um critério ou uma regra de decisão é a especificação de um procedimento que pode ser usado para identificar a melhor alternativa em qualquer problema. Todas as regras de decisão são estudadas com a finalidade de apontar uma escolha, chamada, de modo geral, escolha ótima, ou ótimo.

Por fim, o que vem a ser uma boa decisão, ou melhor ainda, uma boa análise da decisão? Uma boa decisão não é aquela que sempre leva a um final com bons resultados, porque podemos tomar decisões consideradas corretas, mas fatores fora do nosso controle são capazes de transformar resultados vistos como bons, em maus. Ou seja, não se pode medir qualidade de uma decisão apenas por seus resultados, porque a qualidade de uma decisão deve ser considerada igual à qualidade da análise feita para tomar aquela decisão. Porém se tivermos um grande número de situações (repetições), com o mesmo procedimento sendo utilizado, então os resultados permitem avaliar a qualidade das decisões.

3.3 A Análise da Decisão na Agricultura

Pastore ⁴, citado por RODRIGUEZ (1987), diz que a agricultura se apresenta como uma das atividades econômicas mais complexas, dado o caráter aleatório dos vários fenômenos climáticos e biológicos e o grande número de variáveis que afetam as oportunidades de comercialização do produto e o lucro do agricultor. Dessa forma, as decisões devem ser tomadas, sem que se conheçam, com exatidão, suas implicações ou consequências.

Visando auxiliar tal processo de decisão recorre-se, constantemente, aos princípios da teoria da decisão. Pressupõe-se que seja possível especificar convenientemente todas as alternativas de ação possíveis, os estados de natureza (principais fatores que, em conjunto, irão influenciar os resultados de se tomar e implementar uma decisão) e suas probabilidades de ocorrência, as consequências de se implementar cada ação para os estados da natureza considerados, e uma função utilidade a ser maximizada.

Dentre os fatores que influenciam tais resultados, podem-se destacar os fenômenos climáticos e as variações nos custos e preços dos produtos, resultantes das condições de mercado em vigor durante o processo produtivo e de comercialização.

Durante o processo de planejamento e tomada de decisão, o agricultor irá se defrontar com a tarefa de precisar antever os estados da natureza que influenciarão cada alternativa para poder, então, avaliar o resultado de sua implementação. Desta forma, as possibilidades de sucesso estarão vinculadas a probabilidades objetivas e subjetivas do indivíduo, onde as primeiras decorrem de sua experiência passada e as últimas são expectativas que expressam a sua capacidade de antecipar o futuro.

Ressalta-se, portanto, a importância de modelos de planejamento da empresa rural que permitam a inclusão de probabilidades expressas pelo agricul-

⁴PASTORE, J. Decisões em condições de incerteza na agricultura. São Paulo, Fac. Economia e Administração da USP, 1974. 27p. (Trabalho para discussão, 9).

tor, visto que assumem papel fundamental quando se tenta levar em consideração a aleatoriedade de determinados coeficientes.

3.4 Abordagem Normativa e Descritiva da Decisão

Classicamente, a teoria da decisão adota dois enfoques: o primeiro normativo e o segundo descritivo. Um modelo considerado normativo determina como os indivíduos devem decidir, enquanto um modelo descritivo prevê como as pessoas realmente decidem.

Algumas vezes, observa-se que há uma diferença entre a visão descritiva, isto é, como as coisas acontecem, e a visão normativa, isto é, como deveria ser o comportamento em circunstâncias ideais. Modelos com poucos parâmetros e alto poder explanatório-preditivo, são os mais desejados. Os dois enfoques buscam identificar as preferências de um indivíduo ou suas escolhas verdadeiras e, subsequentemente, desenvolver um modelo que caracterize esses padrões de comportamento e que possam ser usados para prever preferências ou escolhas ainda não reveladas.

A distinção entre os dois conceitos é suficiente para se admitir que existem, mas quando se discute como as teorias normativas são usadas na teoria da decisão, a distinção fica mais difícil. Dizer que uma teoria obriga a um certo comportamento, é garantir-lhe um poder ditatorial, porém, o papel da análise da decisão é proporcionar um caminho e um conselho, não uma prescrição.

Deve ser lembrado que o resultado de uma análise da decisão não é uma ordem pela qual uma escolha deve ser feita, mas o entendimento do problema, cuja análise traz uma compreensão de como as crenças e preferências se combinam, e a análise da sensibilidade pode ajudar a verificar se este ajuste é adequado. A análise da sensibilidade é útil porque possibilita examinar as escolhas de uma família de modelos, muitos dos quais possuem algo em comum com o tomador de decisões. Se o modelo escolhido é adequado ao uso dos decisores, tudo bem, caso contrário, deve-se pensar um pouco mais, analisar mais profundamente. Em qualquer uma

das situações, o decisor compreenderá melhor o problema que deve resolver (SILVA, 1990).

A programação linear é um procedimento que proporciona, principalmente, respostas normativas. Por normativo aqui, entende-se a estratégia que deve ser adotada por um indivíduo ou empresa, numa área de produção ou outro setor econômico, quando tanto a finalidade ou objetivo como as condições e restrições que envolvem a ação ou escolha devem tomar uma forma particular.

Poder-se-ia ter definido como objetivo a maximização do lucro de uma empresa rural. As condições e restrições que envolvem os planos a serem seguidos serão as técnicas disponíveis, a quantidade e tipo de solo, a quantidade e qualidade do trabalho, os preços, a quantidade de capital e outros recursos. Estabelecidos o objetivo de maximização do lucro e as restrições de acordo com o tipo e quantidade de recursos, uma resposta indicando os produtos a serem produzidos e técnicas a serem usadas, pode ser especificada. Mas a resposta é fornecida somente em termos do objetivo lucro e das restrições de recursos estabelecidas. Não, necessariamente, explica a estratégia que o proprietário realmente toma. O proprietário pode seguir outra estratégia por ter um outro objetivo além da maximização do lucro, por exemplo, incorrer em menos risco.

Em geral, a programação linear não é um instrumento para análises descritivas. Um exemplo de instrumento utilizado em análise descritiva é dado pelo sistema chamado insumo-produto desenvolvido por LEONTIEF (1951), por se tratar de um método que descreve as inter-relações existentes entre diferentes setores da economia ou de uma indústria em um determinado instante. Pode, também, ser usada para prever as variações necessárias na quantidade de produtos de diversos setores da economia ou da indústria que resulta em um conjunto diferente de bens ou que altera as despesas totais de consumo (RODRIGUEZ, 1987).

3.5 Modelos de Decisão

Uma idéia importante é que as consequências de qualquer ato são determinadas não apenas pela ação em si mesma, mas também por fatores externos, não controlados pelo decisor, cujos comportamentos são desconhecidos para ele na hora da decisão.

Um conceito fundamental no entendimento desses passos é o de estado da natureza (ou estado) que é a descrição completa dos fatores externos independentes do decisor. Se o decisor soubesse o comportamento verdadeiro dos fatores externos, poderia prever, com certeza, a consequência de qualquer ação. Embora assumamos que o decisor não conheça qual estado irá ocorrer na hora em que toma a decisão, ele será capaz de descrever o conjunto dos estados possíveis. O estado que realmente ocorre é chamado de estado verdadeiro.

Em um problema, seja Θ o espaço de todos os estados possíveis θ e \mathcal{A} o espaço de todas as alternativas viáveis a . Por simplicidade, assume-se que estes estados são em número finito, e representados por $\theta_1, \theta_2, \dots, \theta_p$. Um e apenas um destes estados pode ocorrer. Do mesmo modo, assume-se que o número de alternativas também é finito: a_1, a_2, \dots, a_m . Uma e somente uma destas alternativas pode ser escolhida.

Deve-se também quantificar os resultados das ações a serem tomadas, definindo-se uma função lucro $Z(a, \theta)$, normalmente expressa em termos monetários, que especifica o quanto o decisor ganhará ao escolher a alternativa $a \in \mathcal{A}$ se o estado for $\theta \in \Theta$.

Além disso, para cada decisão $a \in \mathcal{A}$, θ possui uma função de densidade de probabilidade $p(\theta|a)$ em $\theta \in \Theta$ com

$$p(\theta|a) \geq 0, \quad \text{para todo } \theta \in \Theta \text{ e } \sum_{\theta \in \Theta} p(\theta|a) = 1$$

Seja x_{ij} a consequência de escolher a alternativa a_i , $1 \leq i \leq m$, quando θ_j , $1 \leq j \leq p$, é o estado verdadeiro. Deste modo, a maioria dos problemas pode ser representada na forma de uma tabela, chamada tabela de decisão (Tabela

1).

Tabela 1. Forma geral de uma tabela de decisão: x_j é a consequência ao escolher a alternativa a_i no estado verdadeiro θ_j

| Ações (Alternativas viáveis) | Estado da natureza (Externos, independentes do decisor) | | | |
|---------------------------------|--|------------|----------|------------|
| | θ_1 | θ_2 | \dots | θ_p |
| a_1 | x_{11} | x_{12} | \dots | x_{1p} |
| a_2 | x_{21} | x_{22} | \dots | x_{2p} |
| \vdots | \vdots | \vdots | \vdots | \vdots |
| a_m | x_{m1} | x_{m2} | \dots | x_{mp} |

Embora os x_{ij} sejam adequados para uma descrição completa das consequências, eles não são necessariamente números mas, assume-se que o tomador de decisão consiga valorizá-los, isto é, ele mede o valor de x_{ij} por uma função real $v(\cdot)$.

Para essa medida de valor deve-se verificar que

$$v(x_{ij}) > v(x_{kl})$$

se e somente se o decisor preferir a consequência x_{ij} à consequência x_{kl} , para $i = 1, 2, \dots, k, \dots, p$ e $j = 1, 2, \dots, l, \dots, m$. Fazendo

$$v_{ij} = v(x_{ij})$$

a forma geral da tabela de decisão passa a ser a Tabela 2.

Tabela 2. Forma geral de uma tabela de decisão quando as consequências x_{ij} são substituídas por seus valores v_{ij}

| Ações (Alternativas viáveis) | Estado da natureza (Externos, independentes do decisor) | | | |
|------------------------------------|--|------------|----------|------------|
| | θ_1 | θ_2 | \dots | θ_p |
| a_1 | v_{11} | v_{12} | \dots | v_{1p} |
| a_2 | v_{21} | v_{22} | \dots | v_{2p} |
| \vdots | \vdots | \vdots | \vdots | \vdots |
| a_m | v_{m1} | v_{m2} | \dots | v_{mp} |

Convém distinguir melhor o significado das palavras risco e incerteza. Situação de risco é aquela na qual se conhecem as probabilidades dos estados da natureza, ao passo que na incerteza, essas probabilidades são desconhecidas. Com base nessas definições os problemas de decisão são agrupados em categorias, de acordo com o conhecimento dos estados da natureza por parte do decisor:

1. **Decisões sob certeza** - assume-se que o verdadeiro estado da natureza, θ_j , $1 \leq j \leq p$, é conhecido pelo decisor antes dele fazer sua escolha, conseguindo prever sem erro os resultados de sua ação.
2. **Decisões sob completa incerteza** - aqui o decisor não sabe nada a respeito dos estados da natureza. Ele não apenas desconhece as probabilidades de ocorrência dos estados, como também não pode quantificar a sua incerteza de modo algum.
3. **Decisões sob risco** - o decisor pode quantificar através de uma distribuição de probabilidades a incerteza da ocorrência dos estados da natureza SILVA (1990).

3.5.1 Critério de decisão em caso de certeza

Neste caso, assume-se que o estado verdadeiro é conhecido pelo decisor antes da sua escolha, de tal sorte que se pode predizer com certeza, as consequências. Ou seja, é o caso em que $p = 1$. A tabela de decisão no caso do estado verdadeiro ser θ_1 , possui a forma trivial da Tabela 3.

Tabela 3. Tabela de decisão sob certeza

| Ações | Estado da natureza |
|----------|--------------------|
| | θ_1 |
| a_1 | v_{11} |
| a_2 | v_{21} |
| \vdots | \vdots |
| a_m | v_{m1} |

Tendo em vista que v_{i1} indica o valor da consequência i quando o estado verdadeiro é θ_1 , e admitindo-se que os valores maiores são preferidos aos menores, a escolha ótima é tomar a ação com o maior valor numérico de v_{i1} .

3.5.2 Critérios de decisão sob estrita incerteza

Alguns critérios comumente usados para escolher uma alternativa em uma situação de completa incerteza, são os seguintes:

Retorno maximin de Wald

Sob a alternativa a_i , a pior consequência que pode ocorrer possui um valor para o decisor de

$$s_i = \min_{j=1}^p [v_{ij}] \quad (4)$$

O s_i é chamado nível de segurança de a_i . Se v_{ij} for interpretado como o valor da consequência x_{ij} , então s_i pode ser interpretado da seguinte forma: a alternativa a_i garante um retorno de ao menos s_i . Wald (1950) sugeriu que o tomador de decisão deveria escolher a_k , tal que ele tivesse o maior nível de segurança possível. Desta forma, o critério maximin de Wald é o seguinte: escolhe-se a_k tal que

$$s_k = \max_{i=1}^m [s_i] = \max_{i=1}^m [\min_{j=1}^p [v_{ij}]] \quad (5)$$

Deve-se notar que este é um critério muito pessimista, porque sua filosofia assume que sempre ocorrerá o pior resultado. Em resumo, para cada alternativa, determina-se o pior resultado: o critério maximin escolhe a alternativa com o “melhor” dos piores resultados.

Índice de otimismo - pessimismo de Hurwicz

De maneira semelhante ao pessimismo de Wald, poder-se-ia ter um critério otimista, considerando-se o melhor resultado de cada ação. O nível otimista da alternativa a_i seria definido como

$$o_i = \max_{j=1}^p [v_{ij}] \quad (6)$$

Desta forma, o_i é o melhor resultado se a_i for escolhida. O critério chamado maximax é o seguinte: escolhe-se a_k tal que

$$o_k = \max_{i=1}^m [o_i] = \max_{i=1}^m [\max_{j=1}^p [v_{ij}]] \quad (7)$$

Em palavras: para cada alternativa, determina-se o melhor resultado; o critério maximax escolhe a alternativa com o “melhor” dos melhores resultados.

Embora existam muitos defensores do critério de Wald, há poucos, se algum, do critério maximax, apesar de haver razão, em princípio, para que se

considere o pessimismo mais racional do que o otimismo. Viu-se que o critério de Wald é o mais cauteloso possível, mas deve-se lembrar que só se ganha alguma coisa se houver alguma aventura.

Em 1951, Hurwicz sugeriu que poucas pessoas desejam ser extremamente tão pessimistas ou tão otimistas quanto as regras os levariam a ser, e propôs um meio termo. Especificamente, argumentou que um tomador de decisão deve ordenar as suas ações de acordo com uma média ponderada dos níveis de segurança e de otimismo.

$$\alpha s_i + (1 - \alpha)o_i, \quad 0 \leq \alpha \leq 1, \quad (8)$$

onde α é o chamado índice de otimismo-pessimismo, um número específico para cada decisor e aplicável a qualquer problema de decisão. A regra de decisão de Hurwicz é a seguinte: escolhe-se a_k tal que

$$[\alpha s_k + (1 - \alpha)o_k] = \max_{i=1}^m [\alpha s_i + (1 - \alpha)o_i] \quad (9)$$

restando apenas estimar α para cada indivíduo, o que pode ser o ponto falho, como assinala McINERNEY (1969), pela dificuldade de se determinar esse valor.

Arrependimento minimax de Savage

Em 1951, Savage argumentou que, pelo uso dos valores v_{ij} para orientar a escolha de alternativas, o decisor está realmente comparando o valor da consequência de uma ação sob um estado da natureza com os valores de todos os outros resultados, quaisquer que sejam os estados da natureza em que eles ocorrem. Mas o estado-verdade está além do controle do tomador de decisão, e assim, a consequência de uma ação deve somente ser comparada com as consequências de outras ações sob o mesmo estado da natureza. Uma particular consequência x_{ij} pode ser ruim no contexto da tabela de decisão, mas pode, apesar disso, ser o me-

lhor resultado de todas as ações se θ_j for o estado-verdade. Então Savage definiu o arrependimento de uma consequência como sendo

$$r_{ij} = (\max_{i=1}^m (v_{ij}) - v_{ij}) \quad (10)$$

ou seja, é igual à diferença entre o resultado da melhor ação, dado que θ_j é o estado-verdade, e o resultado de a_i sob θ_j . Ele argumentou que r_{ij} deve substituir o v_{ij} na tabela de decisão e que, baseado nesta nova tabela, chamada de tabela de arrependimento, o tomador de decisão deve escolher segundo o critério pessimista de Wald, mas lembrando-se que os arrependimentos são ‘perdas’ e não ‘ganhos’. Desta forma, Savage recomenda que à ação seja dado um índice.

$$\rho_i = \max_{j=1}^p [r_{ij}] \quad (11)$$

que é o pior arrependimento que pode resultar da ação a_i , e então a alternativa a ser escolhida deve minimizar ρ_i , ou seja, a escolha a_k tal que

$$\rho_k = \min_{i=1}^m [\rho_i] = \min_{i=1}^m [\max_{j=1}^p [r_{ij}]] \quad (12)$$

Princípios da razão insuficiente de Laplace

Em 1825, Laplace citado por SILVA (1990) argumentou que saber nada acerca do estado-verdade é equivalente a considerar que todos os estados tm a mesma probabilidade de ocorrência. Então o decisor deve escolher a alternativa que maximize o valor médio, ou seja, escolhe-se a_k tal que

$$\sum_{j=1}^p \frac{1}{p} v_{kj} = \max_{i=1}^m [\sum_{j=1}^p \frac{1}{p} v_{ij}] \quad (13)$$

Considerados individualmente, cada um destes critérios parece sensato, mas há situações nas quais levam a decisões diferentes. Então necessita-se

discutir os critérios em profundidade, apresentando razoáveis propriedades de consistência que eles obedeçam em todas as circunstâncias e, desta forma, possa-se determinar qual o melhor critério. Prova-se que é impossível haver critério que obedeça a todas as propriedades e então os problemas sob estrita incerteza são considerados mal-definidos.

Savage ⁵, citado por SILVA (1990), renunciou ao uso das idéias de arrependimento minimax.

3.5.3 Critérios de decisão sob risco

Nas decisões sob risco, o tomador de decisão deve escolher uma decisão $a \in \mathcal{A}$ que maximize seu lucro esperado, a esperança sendo tomada em todo o espaço de resultados Θ . Pela definição de esperança matemática, para seguir tal algoritmo o tomador de decisões maximiza, sob $a \in \mathcal{A}$,

$$L(a) = \sum_{\theta \in \Theta} L(a, \theta)p(\theta, a), \quad (14)$$

onde $L(a)$ denota seu lucro esperado quando ele escolhe a alternativa $a \in \mathcal{A}$. Este é o chamado critério do valor esperado, no qual o decisor escolhe a alternativa que possui o maior valor esperado dos resultados.

Uma decisão que maximiza $L(a)$ em $a \in \mathcal{A}$ comumente é chamada decisão bayesiana. Uma decisão bayesiana será claramente ótima se o algoritmo do máximo valor esperado relacionar-se ao verdadeiro objetivo do decisor.

No dia-a-dia, conhecemos pessoas que possuem diferentes atitudes em relação a situações de risco. Algumas são aversas ao risco, continuamente pagando altos prêmios para se sentirem seguras contra possibilidades que, embora potencialmente desastrosas, possuem muito pouca chance de ocorrer. Outras, as

⁵SAVAGE, L.J. The foundations of statistics 2^a ed., New York: Dover, 1972.

propensas ao risco, raramente fazem seguro, mesmo com uma chance maior de desastres. Convém acrescentar que estas atitudes mudam com as circunstâncias existentes e as importâncias monetárias envolvidas, como, por exemplo, o caso em que uma pessoa muito rica, provavelmente, recusará um seguro de menor valor e enfrentará o risco ela mesma. O mesmo não se pode dizer de uma pessoa com menores recursos.

4 MODELOS COM INCLUSÃO DE INCERTEZA OU RISCO EM PROBLEMAS DE PROGRAMAÇÃO LINEAR

Ao se incluir risco em modelos de programação linear surge o problema de se representar corretamente a aleatoriedade dos coeficientes c_j , a_{ij} e b_i em (1),(2) e (3).

As abordagens para solucionar esse problema podem ser divididas em dois grupos, de acordo com a forma como é tratada a fonte de incerteza. No primeiro grupo, a única incerteza é relacionada à margem bruta c_j por unidade de cada alternativa de produção. Agregam, portanto, as fontes primárias de incerteza (aleatoriedade de margens brutas, preços ou custos) em um único componente de risco expresso na função objetivo dos modelos e que passa, então, a ser o alvo do processo de otimização da programação matemática. No segundo grupo estão as abordagens que, reconhecendo a importância de se considerar a aleatoriedade dos coeficientes técnicos a_{ij} e níveis disponíveis dos recursos na propriedade b_i , incluem as fontes de incerteza ou risco no conjunto de restrições dos modelos.

RODRIGUEZ (1987) faz uma boa revisão dos modelos de inclusão de incerteza , citando:

1) **Minimização da Variância da Margem Bruta Total.** Este método é devido a MARKOWITZ (1952 e 1959) e que HAZELL (1971) afirmou ser a programação quadrática é um método eficiente de se considerar a aleatoriedade da margem bruta total no planejamento agrícola.

2) Miminização dos Desvios Absolutos Totais da Margem Bruta-MOTAD (Minimization of Total Absolute Deviations). Proposto por HAZELL (1971), a partir de modificações introduzidas no método apresentado por MARKOWITZ (1952 e 1959). A modificação sugerida apresenta a vantagem de permitir o uso da programação linear ao substituir a variância pelo desvio absoluto com relação à média como parâmetro de decisão do indivíduo.

3) Modelo com Base no Critério Maximin de Wald. Esse critério apresenta a seguinte formulação. Maximizar

$$V$$

Sujeito a:

$$\sum_{j=1}^n c_{jt}x_j - V \geq 0 \quad (t = 1, \dots, p) \quad (15)$$

$$\sum_{j=1}^n x_j \leq L \quad (16)$$

$$\sum_{j=1}^n a_{ij}x_j \leq \geq b_i, \quad (i = 1, 2, \dots, m) \quad (17)$$

$$x_j, V \geq 0 \quad (j = 1, \dots, n) \quad (18)$$

onde,

V - A menor margem bruta total para todos os p estados da natureza;

c_{jt} - Margem bruta por unidade de área resultante da adoção da atividade A_j dado a ocorrência do estado da natureza S_t ;

x_j - Área total disponível na propriedade agrícola.

a_{ij} - Coeficiente técnico da atividade agrícola A_j na i -ésima restrição.

b_i - Constante, representando a limitação ou disponibilidade do i -ésimo recurso.

Segundo SOARES (1977), este critério pode ser considerado pessimista do ponto de vista do agricultor, uma vez que atribui uma probabilidade igual

a 1 ao pior resultado de cada uma de suas estratégias e uma probabilidade igual a 0 a todos os outros resultados dessa estratégia. Ressalta que sua aplicação seria realista se aplicada a um agricultor de subsistência sem qualquer possibilidade de assumir risco, mas lembra, citando AGRAWAL & HEADY (1972), que a natureza não se comporta como um opositor que, deliberadamente, escolhe sempre a estratégia mais desfavorável ao seu adversário.

4) **Modelo com Base no Critério de Laplace.** Esse critério considera igual a probabilidade de ocorrer qualquer dos estados da natureza. Conforme sugerido por McINERNEY (1969), o método de Laplace está implícito na maioria dos trabalhos de planejamento agrícola que ao utilizarem a programação linear, maximizam o valor esperado da margem bruta total a partir das margens brutas médias esperadas para cada alternativa considerada no modelo (ver Princípios da razão insuficiente de Laplace, p.27).

5) **Modelo com Base no Critério de Hurwicz-** esse modelo pode ser representado da seguinte forma:

Maximizar

$$\sum_{j=1}^n \alpha_j^* x_j \quad (19)$$

Sujeito a:

$$\sum_{j=1}^n a_{i,j} x_j \leq = \geq b_i, \quad (i = 1, \dots, m) \quad (20)$$

$$x_j \geq 0, \quad (j = 1, \dots, n) \quad (21)$$

onde

$$\alpha_j^* = \alpha m_j + (1 - \alpha) M_j \quad (22)$$

em que,

α - medida do pessimismo do agricultor;

$(1 - \alpha)$ - medida do otimismo do agricultor;

m_j - menor margem bruta associada à alternativa A_j ;

M_j - maior margem bruta associada à alternativa A_j .

Segundo McINERNEY (1969), este modelo é visto como um modelo de obtenção de expectativas de resultados incertos, baseados nas atitudes pessimistas ou otimistas do agricultor.

4.1 Modelo com Base no Critério Minimax de Savage

A adoção do critério Minimax de decisão implica na pressuposição de que o agricultor sente insatisfação igual à diferença entre o retorno realmente alcançado e o (maior) retorno que ele poderia ter alcançado se tivesse corretamente previsto o estado da natureza que realmente prevaleceu. Seu objetivo será minimizar o pior resultado previsível. Exige a reformulação da matriz de pagamentos, substituindo cada margem bruta pela diferença que a separa da maior margem bruta existente na coluna. Essa nova matriz poderia ser denominada matriz de desapontamentos (“regret matrix”) e assumiria a forma semelhante a (10).

$$r_{jt} = (\max_t(c_{jt}) - c_{jt}), \quad (23)$$

O plano ótimo será aquele de menor resultado geral negativo (“regret”) máximo possível (V^*), e a formulação do modelo pode ser definido como:

Minimizar

$$V^* \quad (24)$$

Sujeito a:

$$\left(\sum_{j=1}^n r_{jt}x_j\right) - V^* \leq 0, \quad (t = 1, \dots, p) \quad (25)$$

$$\sum_{j=1}^n x_j = L \quad (26)$$

$$\sum_{j=1}^n a_{ij}x_j \leq \geq b_i, \quad (i = 1, \dots, m) \quad (27)$$

$$x_j \geq 0, \quad (j = 1, \dots, n) \quad (28)$$

com todos os coeficientes já definidos anteriormente.

Nota-se, na equação (26), uma limitação ao uso deste modelo. Essa restrição impõe o uso total da área agrícola disponível (L), o que pode não ser vantajoso em alguns casos, e não pode ser alterada para uma inequação do tipo \leq , pois permitiria à função objetivo alcançar o seu mínimo através do não uso dos recursos. HAZZEL (1970) acrescenta também uma crítica ao uso desta abordagem ao verificar que V^* é uma função monotônica crescente de L , o que significa que a solução para diferentes valores de L nunca identificará um valor de V^* representativo de um ponto de mínimo global.

4.2 Maximização da Margem Bruta Total com Restrições de Máxima Perda Admissível

BOUSSARD & PETIT (1967) desenvolveram um modelo de planejamento, utilizando-se de técnicas de programação linear, adequado a situações onde o objetivo é obter um plano que, além de maximizar os ganhos esperados, resulte em uma probabilidade mínima de ruína. Seus autores ressaltam o fato de sua aplicabilidade restringir-se a situações bastantes específicas e de não se tratar de um método que se baseie em alguma medida de utilidade.

O modelo poderia ser assim definido:

Maximizar

$$\sum_{j=1}^n c_j x_j = Z \quad (29)$$

Sujeito a:

$$\left(\sum_{j=1}^n c_j x_j\right) - PTA - STBO = RMA \quad (30)$$

$$\sum_{j=1}^n \bar{c}_j x_j - STBO \leq FUND \quad (31)$$

$$P_j x_j - \frac{1}{k} PTA \leq 0, \quad (j = 1, \dots, n) \quad (32)$$

$$\sum a_{ij} x_j \leq \geq b_i, \quad (i = 1, \dots, m) \quad (33)$$

$$x_j \geq 0, \quad (j = 1, \dots, n) \quad (34)$$

onde,

PTA - perda (prejuízo) total admissível;

RMA - receita mínima admissível;

STBO - capital disponível por empréstimo;

\bar{c}_j - capital requerido para a *j*-ésima alternativa agrícola por unidade área.

P_j - perda (prejuízo), por menor que seja a sua probabilidade de ocorrência, atribuída à *j*-ésima alternativa agrícola por unidade de área; para todos os demais coeficientes valem as definições já dadas anteriormente.

O modelo atribui a denominação de perda total admissível à diferença entre a margem bruta total obtida com o plano ótimo e a receita mínima admissível para subsistência e manutenção do agricultor no ramo. Estabelece também que sempre existirá uma perda P_j , resultado da diferença entre a mais provável e a pior margem bruta atribuível a cada atividade agrícola, nunca superior a uma fração $\frac{1}{k}$ da perda total admissível conforme equação (31).

Desta forma, pode-se verificar que o conjunto de equações definidas pelas expressões (30) e (31) atuam no modelo forçando a definição de um resultado que, além de maximizar a função objetivo, apresente a menor perda total admissível.

4.3 Inclusão de Risco através da Programação Estocástica

A solução ótima de um problema de Programação Linear depende dos parâmetros do mesmo. Uma determinada solução ótima pode permanecer ótima mesmo que alguns dos parâmetros do problema sejam um tanto modificados. Isto, todavia, é insuficiente para garantir que uma solução ótima, para um determinado conjunto de coeficientes, permaneça inalterada caso estes coeficientes, por qualquer razão, se modifiquem após a solução ter sido adotada. Tais modificações não se constituem em eventos raros; ao contrário, frequentemente somos obrigados a planejar e resolver problemas nos quais os coeficientes são apenas estimativas, mais ou menos confiáveis, de valores que se caracterizam por uma distribuição de probabilidade.

Se estas distribuições apresentam uma dispersão pequena em torno das estimativas adotadas, possivelmente a solução ótima computada no modelo determinístico apresentará um desempenho comparável à verdadeira solução ótima (isto é, aquela solução ótima que só pode ser computada a posteriori, quando os elementos probabilísticos já tiverem assumido a característica de evento ocorrido).

Em outros problemas, entretanto, a simples substituição de elementos probabilísticos do modelo por estimativas numéricas dos mesmos, pode levar a decisões cujo desempenho, sob eventos alternativos é simplesmente medíocre, inaceitável quando se considera a experiência de um especialista. Nestes casos o em-

prego de modelos mais elaborados, com consideração explícita dos fatores de risco, pode justificar-se sob o ponto de vista econômico. Um conceito importante para esta avaliação é o de “expectância do custo da incerteza” (LANZER, 1982).

HADLEY (1964) classifica os modelos que se utilizam da programação estocástica em problemas de decisão sequencial e não sequencial. São considerados problemas de decisão sequencial aqueles que envolvem a tomada de duas ou mais decisões em diferentes instantes do horizonte de planejamento e que apresentam a propriedade da última decisão poder ser influenciada não somente por decisões prévias, mas também, por alguns parâmetros estocásticos cujos valores devem ser observados antes que essas decisões sejam tomadas. Identificam-se como problemas de decisão não sequencial aqueles que permitem a tomada de decisões em um único momento, sem que decisões e eventos incertos se intercalem.

Devido à natureza biológica do processo de produção agrícola, a maioria dos problemas de planejamento rural envolve, em maior ou menor grau, um sistema sequencial de tomada de decisões.

De acordo com ANDERSON et al. (1977), esses problemas não são de fácil operacionalização devido ao fato de, pelo menos uma das decisões interrelacionadas não poder ser plenamente especificada antes que um ou mais dos coeficientes aleatórios do modelo seja observado. Para que uma variável de escolha x_j entre na solução ótima, é necessária uma análise do problema, incluindo os valores observados dos coeficientes aleatórios. Isto significa que, *ex ante*, o valor ótimo de x_j não será simplesmente um número, mas sim, função de um ou mais coeficientes aleatórios observados antes da tomada de decisão. Nenhum dos métodos de programação atualmente disponíveis é capaz de resolver esse tipo de problema. Para contornar esse problema, desenvolveu-se um método aproximado de resolução, tendo sido denominado programação estocástica discreta (COCKS, 1968).

RAE (1971a) discute a formulação básica dos métodos de programação estocástica discreta, a elaboração das diversas estruturas de informação que irão alimentá-los e sua flexibilidade ao permitir a formulação de funções objetivo

mais adequadas à teoria bayesiana de decisão.

A programação estocástica discreta pressupõe que os coeficientes aleatórios possam ser apresentados como um conjunto discreto de valores com probabilidade de ocorrência conhecida para cada estado da natureza preestabelecido. Apesar de, em geral, permitir a formulação do modelo através de técnicas de programação linear, este método é de aplicabilidade limitada quando o número de coeficientes considerados aleatórios for grande e, para melhor representação da realidade, se considerar um número elevado de estados da natureza. Desta forma, estes modelos podem rapidamente atingir dimensões impossíveis de serem manipuladas (ANDERSON et al., 1977).

4.4 Inclusão de Risco através da Aplicação Simultânea de Técnicas de Simulação e Programação Linear

Considerando a aleatoriedade dos coeficientes de um problema de programação linear aplicada ao planejamento da empresa rural, considere-se a seguinte formulação:

Maximizar

$$\sum_{j=1}^n \tilde{c}_j x_j = Z \quad (35)$$

sujeito a:

$$\sum_{j=1}^n \tilde{a}_{ij} x_j \leq = \geq \tilde{b}_i \quad (i = 1, \dots, m)$$

$$x_j \geq 0 \quad (j = 1, \dots, n)$$

onde,

Z - Margem bruta total;

x_j - Nível da j-ésima atividade agrícola;

\tilde{c}_j - Margem bruta por unidade da j-ésima atividade que pode assumir valores aleatórios \hat{c}_j ;

\tilde{a}_{ij} - Exigência técnica por unidade da j-ésima atividade agrícola quanto à i-ésima restrição ou fator de produção limitado que pode assumir valores aleatórios \hat{a}_{ij} ;

\tilde{b}_i - Nível (disponibilidade) da i-ésima restrição (fator de produção) que pode assumir valores aleatórios \hat{b}_i .

A aplicação de técnicas de simulação em modelos de programação linear consiste na sucessiva resolução do problema determinístico de programação linear, onde, a cada resolução, os coeficientes aleatórios são substituídos por valores obtidos através de métodos adequados de geração de números aleatórios, com base nas respectivas funções de densidade de probabilidade.

CALATRAVA & DOMINGO (1982) dizem que: “A simulação é um conjunto de técnicas que permite imitar o funcionamento de um sistema, enquanto a programação linear formaliza certos sistemas para otimizá-los segundo algum critério; sistemas estes que, em determinados casos, como o de coeficientes aleatórios, podem ser simulados”.

São raros os trabalhos que utilizam este método, talvez, por ser muito exigente em termos de recursos computacionais. Com o recente aumento da facilidade de acesso a esses recursos e com o aprimoramento das técnicas de simulação em curso, acredita-se ser um método que despertará a atenção de um número maior de pesquisadores.

Por se tratar de dois exemplos que se aproximam bastante da metodologia que será proposta neste trabalho, apresentam-se a seguir dois exemplos de aplicação simultânea de simulação e programação linear, sugeridos por CALATRAVA & DOMINGO (1982) e RODRIGUEZ (1987). No primeiro trabalho os autores selecionaram a priori as estratégias (níveis das atividades agrícolas, x_j) de maior interesse e maior factibilidade de forma indireta, mediante um critério de estabelecimento de posturas a priori do agricultor, com respeito ao universo aleatório

de estados da natureza com os quais irá se defrontar a posteriori da estratégia que adotar.

Assumiram-se distribuições normais e independentes para os preços dos produtos agrícolas componentes dos coeficientes na função objetivo (apesar do método permitir que se considerem também aleatórios os outros coeficientes, assim como outras distribuições).

De acordo com a dispersão dos coeficientes em torno da média, medida através do desvio padrão (S), sete estados da natureza (p) foram estabelecidos a priori, refletindo as diferentes posições do empresário frente ao universo de preços aleatórios:

Tabela 4. Diferentes posições do empresário frente ao universo de preços aleatórios segundo o estado da natureza.

| Estado da natureza(k) | Posição do agricultor | Valor do coeficiente |
|-----------------------|-------------------------|----------------------|
| 1 | Muito otimista | Preço médio + S |
| 2 | Medianamente otimista | preço médio + S/2 |
| 3 | Pouco otimista | preço médio + S/4 |
| 4 | Equilibrada ou média | Preço médio |
| 5 | Pouco pessimista | preço médio - S/4 |
| 6 | Medianamente pessimista | preço médio - S/2 |
| 7 | Muito pessimista | preço médio - S |

Para cada k-ésimo estado da natureza ($k = 1, 2, \dots, p$) resolveu-se o problema não estocástico de programação linear, obtendo-se o vetor X_k de níveis das atividades agrícolas e o valor Z_k da margem bruta total.

Sendo Z_k , para o caso estudado, uma variável com distribuição normal, estimaram-se através da geração de séries de 100 valores para cada um dos preços a partir das distribuições normais adotadas, as respectivas médias, desvios-

padrões e probabilidades de intervalos de classe.

A escolha de um único vetor X_k foi orientada pelo estabelecimento de um intervalo de margens brutas total que satisfizesse ao agricultor e de um valor mínimo abaixo do qual o agricultor se depararia com uma situação “ruinosa”. Para esses intervalos, selecionou-se a estratégia que apresentava a maior probabilidade para o intervalo de margem satisfatória e a menor probabilidade para o intervalo de margens “ruinosas”.

A principal limitação desse trabalho é não apresentar uma alternativa metodológica de inclusão da aleatoriedade de coeficientes correlacionados entre si. Nesse trabalho os autores apenas consideraram distribuições normais para os coeficientes aleatórios.

No segundo trabalho mencionado, RODRIGUES (1986) desenvolveu um método baseado no uso de técnicas de simulação e geração de valores amostrais para o conjunto de coeficientes aleatórios que desempenham papel relevante no problema de programação linear.

O autor discute a necessidade de informações disponíveis para o estabelecimento de funções de densidade de probabilidade adequadas a cada coeficiente aleatório, e as vantagens de se utilizarem as funções normal, triangular e uniforme.

O processo de simulação proposto envolve os três modelos de distribuição referidos anteriormente, mais a eventual existência de correlação entre os coeficientes.

Na geração de valores amostrais para os coeficientes com distribuição normal, considerou-os não correlacionados e correlacionados com distribuição multinormal. Com distribuição triangular, considerou-os não correlacionados e com correlação aproximadamente unitária negativa e positiva. Com distribuição uniforme, considerou-os não correlacionados e com correlação aproximadamente unitária negativa e positiva.

O objetivo do problema de planejamento formulado através da pro-

gramação linear foi maximizar a margem bruta total. Sendo W o conjunto de n coeficientes estocásticos existentes no problema, o processo de simulação teve início com a geração de um primeiro conjunto de valores amostrais. A solução deste problema, incluindo o conjunto w de valores amostrais, aponta para um determinado plano ótimo identificado pelos níveis das variáveis x_j do problema. Este processo foi repetido 1000 vezes. Quando o conjunto de planos ótimos obtido nas primeiras 500 simulações era acrescido de novos planos após a realização das 500 simulações restantes, uma nova série de 500 simulações era implementada até que nenhum plano ótimo novo surgisse como solução no processo de otimização.

De posse dos conjuntos w de valores amostrais gerados e dos planos ótimos obtidos, o próximo passo analisava a distribuição da variável margem bruta total de cada plano ótimo.

Apenas os níveis das variáveis x_j em cada plano ótimo e os coeficientes estocásticos e não estocásticos da função objetivo c_j foram utilizados nesta fase. Se a função objetivo apresentasse coeficientes estocásticos incluídos no conjunto w , cada plano ótimo apresentava margem bruta total variável de acordo com as distribuições individuais de cada coeficiente.

A seleção de um único plano ótimo era orientada através da análise dos parâmetros assimetria, média, moda, variância e probabilidades de ocorrência de valores da margem bruta total superiores a determinados níveis mínimos em cada plano.

A principal limitação desse trabalho é não apresentar uma alternativa metodológica de inclusão da aleatoriedade de coeficientes dependentes entre si no intervalo $[-1, 1]$. O autor considera apenas as distribuições normal, triangular e uniforme para os coeficientes aleatórios.

5 GERAÇÃO DE VARIÁVEIS ALEATÓRIAS MULTIVARIADAS

5.1 Métodos de Geração

A geração em computadores de observações originadas de distribuições multivariadas é mais complexa que de distribuições univariadas, em parte porque a estrutura de dependência multivariada força gerar várias variáveis coletivamente ao invés de individualmente. Em geral, pode-se esperar que o algoritmo e a complexidade computacional seja aumentada, algumas vezes com aumento na dimensão. Em adição, as técnicas univariadas nem sempre generalizam-se prontamente para o caso multivariado. Por exemplo, não existe analogia óbvia do método da função de distribuição inversa.

BOSWELL et al. (1993) informa que a geração multivariada é discutida por JOHNSON (1987), Nachtsheim & Johnson ⁶, Schmeiser ⁷ e Schmeiser & Lal ⁸ Barone ⁹ considera a geração de séries temporais multivariadas. Estudo de simulação algumas vezes empregando técnicas da redução da variância que requer a geração de variáveis aleatórias correlacionadas pode ser visto, por exemplo, em

⁶NACHTSHEIM, C.J. & JOHNSON, E. A new family of multivariate distributions with applications to Monte Carlo Studies. **J. Amer. Statist. Assoc.** 83. 984-89. 1988.

⁷SCHMEISER, B.W. Simulation experiments In: D.P. HEYMAN & M.J. SOBEL. eds. **Stochastic Models**. Elsevier: Amsterdam. 1990. Chap.7. p.295-330.

⁸SCHMEISER, B.W. & LAL, R. Squeeze methods for generating gamma variates. **J. Amer. Statist. Assoc.** n.75, p.679-682, 1980.

⁹BARONE, P. A method of generating independent realizations of a multivariate normal stationary and invertible ARMA (p,q.) process. **J. Time Ser. Anal.** n.8, p.125-30, 1987.

Hammersley & Handscomb ¹⁰, Kleijnen ¹¹, Kleijnen ¹², e Sim e Lee ¹³.

Existem três importantes métodos para gerar variáveis aleatórias multivariadas. São eles: (i) o método das distribuições condicionais, (ii) o método da transformação, e (iii) método da rejeição, JOHNSON (1987) e BOSWELL et al. (1993). Descrever-se-á a seguir sucintamente os métodos (ii) e (iii), ficando a descrição do método (i) para mais adiante, uma vez que foi o método utilizado nesse trabalho.

Método da transformação

Se \mathbf{X} e \mathbf{Y} são vetores aleatórios satisfazendo uma relação funcional

$$\mathbf{Y} = h(\mathbf{X})$$

então as observações de \mathbf{Y} podem ser geradas, gerando primeiro as observações de \mathbf{X} e então aplicando a transformação $\mathbf{Y} = h(\mathbf{X})$. Essas transformações podem ser lineares ou não.

Uma importante aplicação de transformações lineares é na geração de observações multivariadas cujo vetor de médias $\boldsymbol{\mu}$ e matriz de dispersão $\boldsymbol{\Sigma}$ são especificadas, mas cuja distribuição é desconhecida ou não especificada. Tais observações são utilizadas, por exemplo, no estudo das propriedades de mínimos

¹⁰HAMMERSLEY, J.M. & HANDSCOMB, D.C. Monte Carlo Methods. Methuen. London. 1964.

¹¹KLEIJNEN, J.P.C. Estatistical techniques in simulation. Part 1. Marcel Dekker. New York. 1974.

¹²KLEIJNEN, J.P.C. Estatistical techniques in simulation. Part 2. Marcel Dekker. New York. 1975.

¹³SIM, C.H. & LEE, P.A. Simulation of negative binomial processes. **J. Statist. comput. simulation** n.34. p.20-42. 1989.

quadrados, incluindo o “kriging” de dados espacial Alabert ¹⁴ e Davis ¹⁵. Suponha que o vetor aleatório k dimensional \mathbf{Y} é especificado com média $\boldsymbol{\mu}$ e matriz de dispersão $\boldsymbol{\Sigma}$. Por simplicidade, toma-se $\boldsymbol{\Sigma}$ como definida positiva. De acordo com a decomposição de Cholesky, $\boldsymbol{\Sigma}$ pode ser escrita como

$$\boldsymbol{\Sigma} = \mathbf{H}\mathbf{H}'$$

onde \mathbf{H} é uma matriz triangular inferior k por k ; por exemplo, veja-se George e Liu (1981)¹⁶ ou Golub ¹⁷. Agora seja \mathbf{X} um vetor aleatório k dimensional cujas variáveis aleatórias são independentes com média 0 e variância unitária. Então

$$E(\mathbf{X}) = \mathbf{0}$$

e

$$E(\mathbf{X}\mathbf{X}') = \mathbf{I}$$

da qual é visto que a transformação

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \boldsymbol{\mu}$$

gerará \mathbf{Y} com os momentos desejados. Certamente, \mathbf{X} é relativamente fácil de gerar uma vez que suas variáveis aleatórias são independentes.

Como um outro exemplo de aplicação desse método, tem-se a distribuição Poisson multivariada. Se X, X_1, \dots, X_k são variáveis aleatórias de Poisson independentes com respectivas médias $\lambda, \lambda_1, \lambda_2, \dots, \lambda_k$ então o vetor aleatório

¹⁴ALABERT, F. The practice of fast conditional simulations through the LU decomposition of the covariance matrix. *Math. Geol.* n.19. p.369-86. 1987.

¹⁵DAVIS, M.V. Production of conditional simulations via the LU triangular decomposition of the covariance matrix. *Math. Geol.* n.19. p.91-98. 1987.

¹⁶GEORGE, A. & LIU, J.W.H. Computer solution of large sparse positive definitive systems. Prentice-Hall, Englewood Cliffs, NJ. 1981.

¹⁷GOLUB, G.H. Matrix decompositions and statistical calculations. In: R.C. MILTON & J.A. NELDER, ed. Statistical computations. Academic Press. New York. 1969.

$$\mathbf{Y} = (Y_1, \dots, Y_k)'$$

definido por

$$Y_i = X + X_i, \quad i = 1, \dots, k$$

é dito ter a distribuição de Poisson multivariada com parâmetros $\lambda, \lambda_1, \lambda_2, \dots, \lambda_k$ (Johnson e Kotz ¹⁸. Para gerar uma observação $Y = y$ da distribuição de Poisson multivariada, pode-se gerar $k + 1$ variáveis Poisson independentes, X, X_1, \dots, X_k , e então fazer a transformação definida anteriormente. Note-se que não há nenhum conhecimento, antes da geração, da estrutura de dependência de Y_i , para todo i .

Método da rejeição

O método da rejeição é muito útil para geração univariada, e é claro que este método não é restrito ao caso de uma dimensão. Em princípio, a rejeição multivariada é uma generalização direta do método de rejeição univariada e é apresentado de forma resumida no seguinte algoritmo. O vetor aleatório de interesse é \mathbf{X} com densidade multivariada f , e \mathbf{Y} com densidade g é um vetor aleatório cujas ocorrências de suas observações são aceitas ou rejeitadas para produzir a desejada ocorrência de \mathbf{X} .

Passo 1. Calcule $m = \sup_x [f(x)/g(x)]$, supremo do cociente.

Passo 2. Gere uma observação y de g .

Passo 3. Gere $U = u$ de uma distribuição uniforme no intervalo $[0;1]$.

Passo 4. Se $u < f(y)/\{mg(y)\}$, então atribua $\mathbf{X} = y$.

Passo 5. Volte ao passo 2.

Entretanto, existem algumas dificuldade práticas desse algoritmo. Uma grande dificuldade está em encontrar a função densidade g que seja próxima

¹⁸JOHNSON, N.L. & KOTZ, S. Distributions in statistics: discrete distributions. Houghton Mifflin, Boston. 1969

a densidade multivariada de interesse. Para simplificar a geração de g no passo 2, é conveniente escolher a distribuição de \mathbf{Y} de modo que seus elementos (variáveis aleatórias) sejam independentemente distribuídos. Uma possível solução é fazer com que \mathbf{Y} tenha elementos independentes Y_i com as mesmas distribuições marginais com o correspondente elemento de \mathbf{X} . Certamente essa solução não é destituída de dificuldades. Primeiro, não existe garantia que a constante m que aparece no passo 1 seja finita. Por outro lado, se a dependência entre os elementos de \mathbf{X} é forte, então não se pode esperar que a densidade seja bem aproximada pela densidade tendo marginais independentes.

5.2 Distribuições Bivariadas Correlacionadas

KLEIJNEN (1974) descreve brevemente, como gerar amostras de duas variáveis correlacionadas. O esquema geral apresentado é o seguinte:

Se X e Y são duas variáveis aleatórias que devem ter um coeficiente de correlação ρ (e devem ter média e variância μ_1, σ_1^2 e μ_2, σ_2^2 , respectivamente), então usando

$$Y = aX + b + U$$

com

$$a = \rho \frac{\sigma_2}{\sigma_1}, \quad b = \mu_2 - \rho \frac{\sigma_2}{\sigma_1} \mu_1$$

$$U = (1 - \rho^2)^{\frac{1}{2}} \sigma_2 Z$$

onde U é independente de X , e Z tem média 0 e variância 1, geram-se duas amostras com correlação ρ .

Uma condição severa é que X e Y devem ter um particular tipo de distribuição (além de satisfazer as condições sobre médias e variâncias). Então

o esquema proposto só funciona se U é selecionado corretamente. Por exemplo, supondo-se que X e Y são variáveis normais, a soma de variáveis normais permanecerá normal, mas essa propriedade não é válida para X e Y com distribuição uniforme (ou triangular) já que a soma de variáveis uniformes não permanece uniforme (ou triangular). Por exemplo, a soma de duas variáveis aleatórias uniformes é triangular e a soma de duas triangulares toma forma setorizada quadrática, tendendo à normal, lembrando o teorema do limite central.

BARNETT (1980) e JOHNSON (1986) afirmam que é comum encontrar dados bivariados em uma região limitada onde as distribuições marginais são uniformes mas existe clara indicação de dependência entre as duas variáveis aleatórias. É fácil construir famílias de distribuições uniformes bivariadas dependentes. Se (X, Y) tem uma distribuição bivariada contínua com funções de distribuição marginais $F_X(x)$ e $F_Y(y)$, respectivamente, é suficiente considerar

$$U = F_X(X), \quad V = F_Y(Y)$$

para que (U, V) seja uniforme bivariada, no sentido de que as variáveis U e V são uniformemente distribuídas no intervalo $[0;1]$. A forma de distribuição de (U, V) derivada de diferentes famílias de distribuição para (X, Y) foi apresentada, por exemplo por Kimeldorf e Sampson (1975) para o caso onde (X, Y) são tipos bivariados de Plackett (1965), Morgenstern (1956), exponencial de Marshall-Olkin (1967), exponencial de Gumbel (1960), Cauchy (Mardia, 1970) e Mardia-Pareto (1962).

Somente amplos resultados quantitativos sobre a natureza da dependência entre (U, V) estão disponíveis. Se (X, Y) são independentes, então (U, V) também são independentes, manifestada em uma função densidade de probabilidade conjunta constante. Qualquer variação no valor da função densidade de probabilidade é indicativo da dependência entre (U, V) . É intuitivamente razoável que a dependência entre X e Y traduza uma medida de dependência entre U e V .

Esse assunto é discutido para diferentes medidas de dependência

por Kimeldorf e Sampson ¹⁹. Para famílias (X, Y) de um parâmetro θ é particularmente interessante examinar a natureza da dependência entre U e V para diferentes valores daquele parâmetro, θ , e de fato para diferentes famílias básicas da distribuição para (X, Y) . Uma modesta tentativa nesta direção é providenciada a seguir onde gráficos de contornos da função densidade de probabilidade de (U, V) são apresentadas para alguns casos representativos. Consideraram-se as distribuições uniformes derivadas da Morgernstern, Plackett, exponencial de Gumbel, normal e Cauchy para (X, Y) .

Morgernstern ²⁰ introduziu uma família uni-paramétrica de distribuições com função de distribuição conjunta

$$F(x, y) = F(x)G(y)\{1 - \psi[1 - F(x)][1 - G(y)]\}, \quad |\psi| < 1$$

onde, $F(x)$ e $G(y)$ são funções de distribuição univariadas contínuas. É particularmente simples transformar para uniforme bivariada (U, V) . Fazendo

$$U = F(X), \quad V = G(Y)$$

encontramos que (U, V) tem função densidade de probabilidade

$$f(u, v) = 1 - \psi(1 - 2u)(1 - 2v)$$

as distribuições marginais de (U, V) são uniformes no intervalo $[0;1]$. Note-se que qualquer que seja a forma explícita de $F(x)$ e $G(y)$ obtemos uma única família uni-paramétrica para (U, V) . Desde que

¹⁹KIMELDORF, G. & SAMPSON, A. One-parameter families of bivariate distributions with fixed marginals. **Commun. Statist.** n.4, p.293-301. 1975.

²⁰MORGERNSTERN, D. Einfache Beispiele Zweidimensionaler Verteilungen. **Mitt. Math. Statist.** 8 234-235. 1956.

$$E(UV) = \frac{1}{4} - \frac{\psi}{(36)}$$

o coeficiente de correlação é

$$\rho_{U,V} = -\frac{\psi}{3}$$

e

$$|\rho_{U,V}| \leq \frac{1}{3}$$

A natureza da dependência entre U e V é ilustrada na Figura 1 onde um gráfico de contornos de $f(u, v)$ é dado para o caso $\psi = 1/2$, ($\rho_{(U,V)} = -1/6$). O que se nota é a densidade unitária constante sobre as linhas $u = 1/2$ e $v = 1/2$ (o caso universal); com suaves depressões nos cantos nordeste e sudeste compensadas nos cantos noroeste e sudoeste.

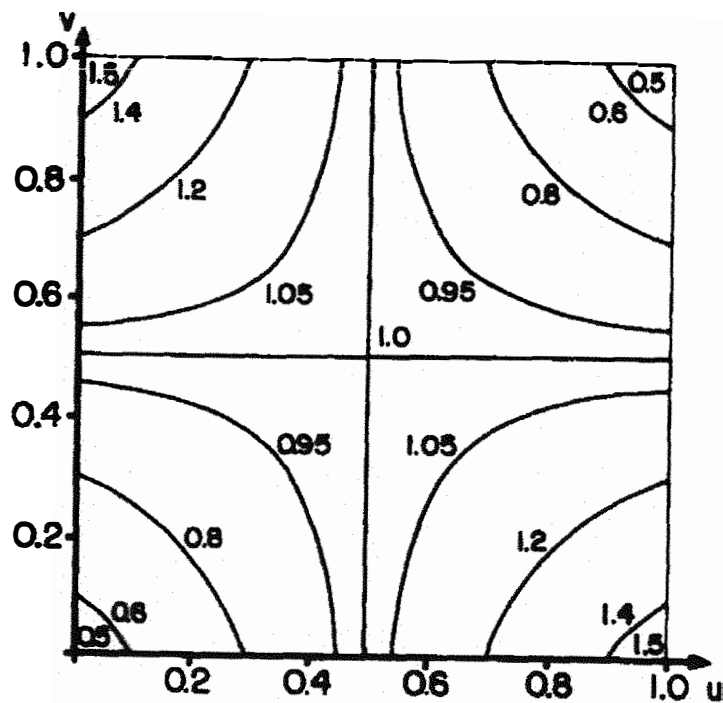


Figura 1 - Contornos da distribuição uniforme bivariada de Morgenstern para $\psi = 1/2$.

Uma outra família de distribuições bivariadas uni-paramétrica devido a Plackett ²¹ é dada implicitamente por

$$\frac{F(x, y)[1 - F_X(x) - F_Y(y) + F(x, y)]}{[F_X(x) - F(x, y)][F_Y(y) - F(x, y)]} = 1 - \alpha \quad (\alpha < 1)$$

onde para quaisquer funções de distribuição marginais $F_X(x)$ e $F_Y(y)$ a função distribuição conjunta é $F(x, y)$. Qualquer que seja a forma de $F_X(x)$ e $F_Y(y)$ a correlação entre X e Y é

$$\rho_{X,Y} = [-\alpha(2 - \alpha) - 2(1 - \alpha)\ln(1 - \alpha)] / \alpha^2$$

Mardia ²². Novamente a transformação para a distribuição uniforme é particularmente direta e independente das formas $F_X(x)$, $F_Y(y)$. Tem-se

$$f(u, v) = \frac{(1 - \alpha)[(2uv - u - v)\alpha + 1]}{[\alpha(u + v) - 1]^2 + 4\alpha(1 - \alpha)uv}$$

Agora, $\rho_{X,Y} > 0$ se $\alpha < 0$ e $\rho_{X,Y} < 0$ se $0 < \alpha < 1$. Também se $|\alpha|$ é pequeno, a família de distribuições bivariadas de Plackett é essencialmente da forma de Morgenstern, Mardia²³.

Assim, para valores intermediários de α espera-se que a forma de $f(u, v)$ exiba alguma das características da Figura 1, mas com aumento da complexidade da estrutura quando $|\alpha|$ cresce. Isto é ilustrado pelo gráfico de contornos da Figura 2 onde $\alpha = -4$ ($\rho_{X,Y} = 0,494$). (Note-se, que certamente tem-se agora associação positiva entre U e V , em contraste com a associação negativa na Figura

²¹PLACKETT, R.L. A class of bivariate distributions. **J. Amer. Statist. Assoc.** n.60, p.516-22. 1965.

²²MARDIA, K.V. Some contributions to contingency-type bivariate distributions. **Biometrika.** n.54, p.235-49. 1967.

²³MARDIA, K.V. The performance of some tests of independence for contingency-type bivariate distributions. **Biometrika.** n.56, p.449-51. 1969.

1). Vê-se que $f(u, v)$ toma os valores finitos $(1 - \alpha)$ e $(1 - \alpha)^{-1}$ nos pares de cantos opostos.

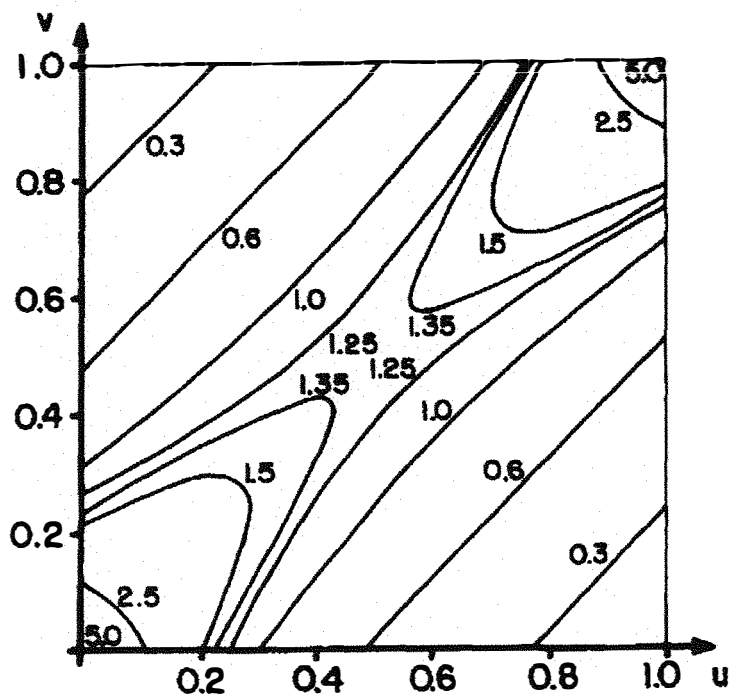


Figura 2 - Contornos da distribuição uniforme bivariada de Plackett : $\alpha = -4$

Outra distribuição é a de Gumbel²⁴. Suponha que (X, Y) tenha uma distribuição exponencial de Gumbel com função densidade de probabilidade

$$f(x, y) = [(1 + \theta x)(1 + \theta y) - \theta] \exp(-x - y - \theta xy), \quad (0 < \theta < 1)$$

Para $\theta = 0$, X e Y são exponenciais padrão independentes. Com $\theta > 0$, X e Y são negativamente correlacionadas com $|\rho_{X,Y}|$ alcançando o seu valor máximo quando $\theta = 1$, onde $\rho_{X,Y} = -0,4036$. Agora obtém-se a distribuição uniforme bivariada com

²⁴GUMBEL, E.J. Bivariate exponential distributions. *J. Amer. Statist. Assoc.* n.55, p.698-

$$f(u, v) = \{(1 - \theta) - \theta[\ln(1 - u) + \ln(1 - v)] + \theta^2 \ln(1 - u) \ln(1 - v)\} \exp[-\theta \ln(1 - u) \ln(1 - v)]$$

É evidente que $f(u, v)$ tem um ponto de sela. Isso ocorre quando $u = v$ com $v = 1 - e^{-\eta}$ onde η é a raiz real da equação cúbica

$$\theta^2 \eta^3 + 2\theta \eta^2 + (1 - 2\theta)\eta - 1 = 0$$

A forma de $f(u, v)$ é calculada para $\theta = 0,5$ ($\rho_{X,Y} = -0,277$).

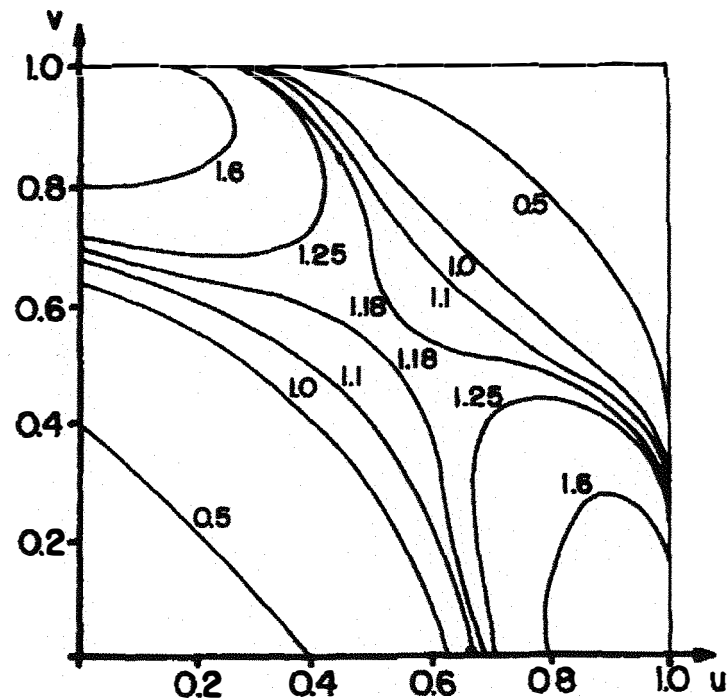


Figura 3 - Contornos da distribuição uniforme bvariada de Gumbel: $\theta = 1/2$.

A Figura 3 apresenta um gráfico de contornos da correspondente $f(u, v)$. O ponto de sela ocorre quando $u = v = 0,552$. A densidade é caída, do valor do ponto de sela de 1.185 a zero no canto $(0,0)$ e $(1,1)$ e sobe indefinida e abruptamente nos cantos $(0,1)$ e $(1,0)$.

Mardia ²⁵ descreve uma distribuição bivariada de Cauchy com função de densidade de probabilidade

$$f(x, y) = \frac{1}{2\pi} (1 + x^2 + y^2)^{-\frac{3}{2}}$$

Espera-se algum comportamento extremo na correspondente uniforme bivariada, cuja função densidade de probabilidade é

$$f(u, v) = \frac{\pi}{2} \frac{[1 + \tan^{-2}(\pi u)][1 + \tan^{-2}(\pi v)]}{[1 + \tan^{-2}(\pi u) + \tan^{-2}(\pi v)]^{\frac{3}{2}}}$$

O gráfico de contornos mostrado na Figura 4 mostra algumas regiões interessantes, com um platô central relativamente amplo, raso e plano, e infinitamente alto terminando em pico em todos os quatro cantos. Nota-se que $f(u, v) = 0$ ao longo das quatro margens, exceto para os cantos extremos onde ela é infinita.

Essas distribuições conjuntas são estranhas, sendo difícil imaginar duas variáveis, na prática que possam obedecê-las. Pressupor que as distribuições marginais são uniformes pode parecer algo perfeitamente razoável à primeira vista. Mas pode ser que o pesquisador deixe de considerar isso razoável quando toma conhecimento da forma estranha que tem a distribuição conjunta.

²⁵MARDIA, K.V. Families of Bivariate Distributions. Darien, Conn. Hafner Publishing Co. 1970.

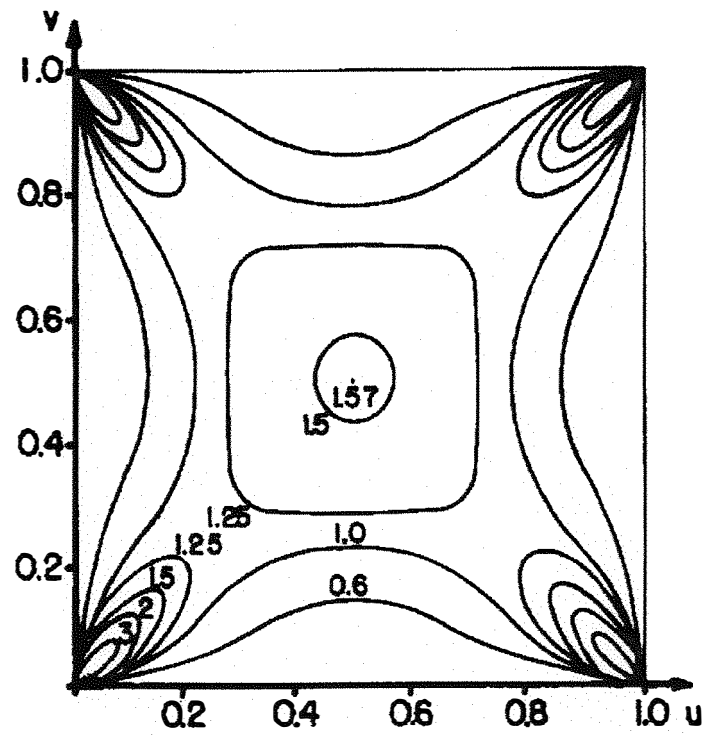


Figura 4 - Contornos da distribuição uniforme bivariada de Cauchy.

6 METODOLOGIA PROPOSTA

O método desenvolvido neste trabalho está baseado no uso de técnicas de simulação e geração de valores amostrais para o conjunto de coeficientes aleatórios que desempenhem papel relevante no problema de programação linear.

De acordo com as informações disponíveis faz-se necessário o estabelecimento de funções de distribuições de probabilidades adequadas a cada coeficiente aleatório.

Uma vez estabelecida uma função de distribuição de probabilidades para cada coeficiente aleatório, apresentam-se a seguir, os métodos de geração de valores amostrais para cada tipo de distribuição.

6.1 Adoção de Funções Probabilísticas para os Coeficientes Aleatórios

Todo coeficiente aleatório pode assumir uma determinada distribuição de probabilidades. Obter o modelo probabilístico mais adequado implica na análise prévia da quantidade e qualidade dos dados disponíveis. Séries suficientemente longas de dados históricos de preços e custos, por exemplo, seriam necessárias para ajustar distribuições teóricas convenientes, tais como a Normal, Beta, Poisson, Gama, Uniforme, Triangular, etc., assim como para se estudar a eventual existência de correlação entre os coeficientes aleatórios.

Quando poucas informações ou apenas probabilidades subjetivas estão disponíveis, sugere-se a utilização de distribuições mais simples que resultem em uma boa aproximação. A distribuição Triangular é uma que se destaca neste

caso, devido à sua simplicidade, capacidade de expressar distribuições assimétricas, e exigência de poucos parâmetros que a descrevam (ANDERSON et. al., 1977).

Na ausência quase total de informações acerca da distribuição de probabilidades do coeficiente aleatório, uma aproximação razoável pode ser, às vezes, obtida ao se utilizar a distribuição Uniforme.

As distribuições Triangular e Uniforme são as que mais facilmente permitem o aproveitamento das opiniões e probabilidades subjetivas expressas pelo agricultor. Em geral, o agricultor é capaz de manifestar, em termos discretos, a sua expectativa acerca da frequência com que determinado coeficiente assume certos valores, ou ainda, valores mínimos, máximos e mais provável que o coeficiente pode assumir. Estes parâmetros são suficientes para o ajuste de uma certa distribuição Triangular. O caso da distribuição Uniforme se ajusta a situações onde o agricultor não é capaz de prever o valor de ocorrência mais frequente, sendo suficiente que informe apenas o intervalo de variação do coeficiente aleatório e que se possa pressupor uma probabilidade uniforme para qualquer valor dentro desse intervalo.

6.2 Uso Simultâneo de Métodos de Simulação e Programação Linear.

O uso da simulação terá como objetivo gerar valores apropriados para os coeficientes aleatórios. O conjunto formado pela geração de um valor para cada coeficiente estocástico, a partir das respectivas distribuições de probabilidade, será incluído no problema de programação linear. A solução deste problema apontará para uma única estratégia agrícola ótima e será função dos valores determinísticos adotados.

Repetindo-se o processo de simulação um número suficientemente grande de vezes, verificar-se-á se os resultados da programação linear compõem um conjunto finito de estratégias agrícolas, e com que frequência cada estratégia ocorre. Desta forma, seriam conhecidos *ex ante* as estratégias que, dependendo do

estado da natureza que prevalecesse *expost*, otimizariam o objetivo estabelecido no planejamento da empresa.

O nível de confiabilidade dos resultados deste processo depende da perfeita representação da aleatoriedade dos coeficientes do problema de programação linear, pois, as estratégias ótimas e respectivas frequências serão reflexo do efeito conjunto das distribuições individuais de cada coeficiente estocástico.

Neste trabalho, foram utilizados os modelos normal, triangular e uniforme para ajuste da distribuição de probabilidade dos coeficientes estocásticos. A escolha do modelo mais adequado dependeu dos dados disponíveis. Existiu uma preocupação em demonstrar a viabilidade de se considerar a eventual correlação existente entre os coeficientes aleatórios que, para definir corretamente cada estado da natureza, devem ser considerados no processo de simulação.

6.3 Métodos de Obtenção de Valores Amostrais para os Coeficientes Aleatórios do Problema de Programação Linear.

O processo de simulação proposto, envolvendo os modelos de distribuição referidos no item anterior, mais a eventual existência de correlação entre os coeficientes, considerará as seguintes situações:

- 1) Geração de valores amostrais para os coeficientes com distribuição normal independentes e correlacionados com distribuição multinormal.

Foi utilizado o método das distribuições condicionais apresentado por GODOI (1978) e BOSWELL et al. (1993). Uma distribuição k -dimensional pode ser representada como um produto de k distribuições condicionais. Correspondentemente, a geração de variável aleatória de uma distribuição multivariada pode ser obtida gerando uma observação de cada uma das distribuições condicionais

univariadas em sequência. A densidade conjunta de um vetor aleatório \mathbf{X} pode ser fatorada como segue:

$$f(x_1, x_2, \dots, x_k) = f_1(x_1)f_2(x_2|x_1)\dots f_k(x_k|x_1, \dots, x_{k-1})$$

O método de geração é descrito pelo seguinte algoritmo:

Suponha que \mathbf{X} segue a distribuição normal multivariada não singular com vetor de médias $\boldsymbol{\mu}$:

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1N} \\ x_{21} & x_{22} & \cdots & x_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ x_{k1} & x_{k2} & \cdots & x_{kN} \end{pmatrix} = (X_1, \dots, X_k)'$$

$$\boldsymbol{\mu} = (\mu_1, \dots, \mu_k)'$$

e matriz de dispersão semidefinida positiva

$$\boldsymbol{\Sigma} = (\sigma_{ij})$$

O i -ésimo elemento da diagonal de $\boldsymbol{\Sigma}$ é

$$\sigma_i^2 = \sigma_{ii}$$

Para $i = 1, \dots, k$, definimos

$$\mathbf{X}_{(i)} = (X_1, \dots, X_i)'$$

como o vetor dos primeiros i componentes de \mathbf{X} . O vetor de médias de $\mathbf{X}_{(i)}$ é,

$$\boldsymbol{\mu}_{(i)} = (\mu_1, \dots, \mu_i)',$$

e a matriz de dispersão de $\mathbf{X}_{(i)}$, dada por

$$\Sigma_{ii} = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1i} \\ \vdots & \ddots & \vdots \\ \sigma_{i1} & \cdots & \sigma_{ii} \end{pmatrix} = \begin{pmatrix} \text{Var}(X_1) & \cdots & \text{Cov}(X_1, X_i) \\ & \ddots & \vdots \\ & & \text{Var}(X_i) \end{pmatrix}$$

é a submatriz i por i de Σ . O vetor de covariância de X_{i+1} com X_1, \dots, X_i é dado por

$$\boldsymbol{\sigma}_{(i)} = (\sigma_{1,i+1}, \dots, \sigma_{i,i+1})$$

A distribuição marginal de X_1 é normal univariada com média μ_1 e variância σ_1^2 ; ao passo que, para $i = 1, \dots, k-1$, a distribuição condicional de X_{i+1} , dado X_1, \dots, X_i é normal univariada com média

$$\bar{\mu}_{i+1} = \mu_{i+1} + \boldsymbol{\sigma}'_{(i)} \Sigma_{ii}^{-1} (\mathbf{X}_{(i)} - \boldsymbol{\mu}_{(i)})$$

e variância

$$\bar{\sigma}_{(i+1)}^2 = \sigma_{i+1}^2 - \boldsymbol{\sigma}'_{(i)} \Sigma_{ii}^{-1} \boldsymbol{\sigma}_{(i)}$$

Isso origina o seguinte algoritmo para gerar a distribuição normal multivariada k -dimensional com média $\boldsymbol{\mu}$ e matriz de dispersão Σ .

Passo 1. Faça $i = 1$, $\boldsymbol{\mu} = \mu_1$, e $\sigma^2 = \sigma_1^2$.

Passo 2. Gere $X_i = x_i$ de uma $N(\boldsymbol{\mu}, \sigma^2)$.

Passo 3. Se $i = k$, então atribua a $\mathbf{X} = (x_1, x_2, \dots, x_k)'$.

Passo 4. Incremente i

Passo 5. Calcule $\boldsymbol{\mu} = \bar{\boldsymbol{\mu}}_i$ e $\sigma^2 = \bar{\sigma}_i^2$

Passo 6. Volte ao passo 2.

2) Geração de valores amostrais para os coeficientes com distribuição uniforme independentes e dependentes.

Nesse caso, os valores (u) são obtidos a partir da geração de valores com distribuição normal e aplicando a função de distribuição normal $F_X(x)$ que tem distribuição uniforme no intervalo $[0;1]$. A mesma correlação com que os valores são gerados para a normal também o são para as variáveis uniformes. Assim, se X_1, X_2 são variáveis aleatórias normais dependentes, com correlação ρ_{X_1, X_2} então

$$F(X_1) = U_1 \text{ e } F(X_2) = U_2$$

são variáveis aleatórias uniformes com correlação

$$\rho_{U_1, U_2} = \frac{6}{\pi} \arcsen\left(\frac{1}{2}\rho_{X_1, X_2}\right)$$

(WILKS, 1962). Essa relação entre ρ_{U_1, U_2} e ρ_{X_1, X_2} se aproxima bem de uma reta passando pela origem no intervalo $[0;1]$, o que pode ser visto pelo gráfico dessa função ou pela inspeção das duas primeiras colunas da Tabela 5, onde a_{ij} corresponde a ρ_{U_1, U_2} e ρ_{ij} à ρ_{X_1, X_2} , ou seja, da relação acima tem-se

$$\rho_{ij} = 2\text{sen}\left(a_{ij}\frac{\pi}{6}\right)$$

Segundo BARNETT (1980) a correspondente variável aleatória uniforme bivariada (U_1, U_2) tem a seguinte função densidade de probabilidade conjunta:

$$f(u_1, u_2) = \frac{1}{\sqrt{1 - \rho_{X_1, X_2}^2}} \exp\left\{-\frac{1}{2}\left\{\frac{[\rho_{X_1, X_2}\phi^{-1}(u_1) - \phi^{-1}(u_2)]^2}{(1 - \rho_{X_1, X_2}^2)} - \phi^{-2}(u_2)\right\}\right\}$$

onde,

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2} dt$$

é a função de distribuição da normal padrão. Aqui, $f(u_1)$ e $f(u_2)$, são funções densidade de probabilidade marginais das variáveis aleatórias, U_1 e U_2 , uniformes no intervalo $[0;1]$.

contornos de $f(u_1, u_2) = c$ (constante) são obtidos de:

$$u_2 = \phi\left(\frac{1}{\rho_{X_1, X_2}}[\phi^{-1}(u_1) \pm \sqrt{\phi^{-2}(u_1)(1 - \rho_{X_1, X_2}^2) + k}]\right)$$

onde

$$k = -2\ln(c\sqrt{1 - \rho^2})$$

Uma forma simples aparece quando

$$c = (1 - \rho_{X_1, X_2}^2)^{-\frac{1}{2}}$$

quando

$$u_2 = \phi\left[\frac{1 \pm \sqrt{1 - \rho_{X_1, X_2}^2}}{\rho_{X_1, X_2}}\phi^{-1}(u_1)\right]$$

que produz contornos em "forma de oito" do canto (0,0) ao oposto (1,1) (para $\rho_{X_1, X_2} > 0$). A $f(u_1, u_2)$ é ilustrada para o caso $\rho_{X_1, X_2} = 0,8$ pelo gráfico de contornos da Figura 5. As encostas nos cantos sudeste e nordeste são infinitamente altas.

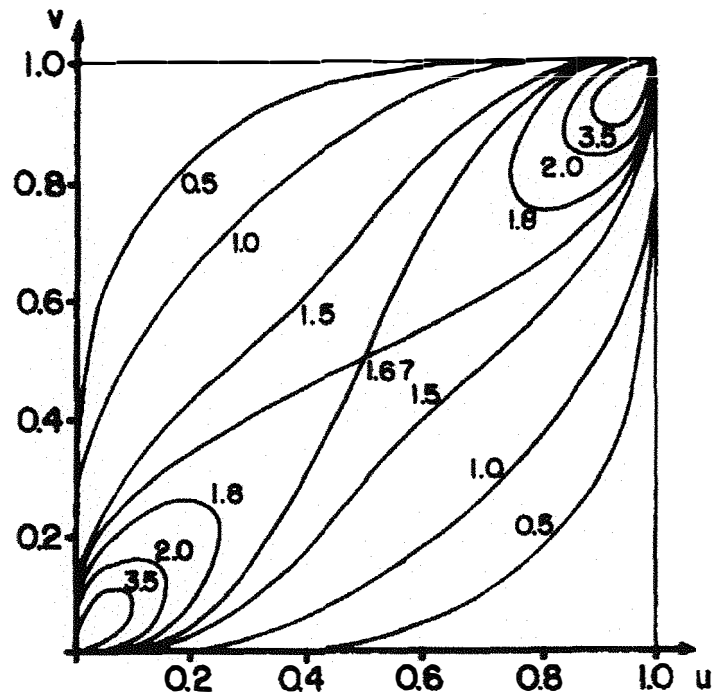


Figura 5. Contornos da distribuição uniforme bivariada originada da normal.

Uma generalização para n variáveis uniformes pode ser facilmente obtida, simplesmente gerando n variáveis normais e aplicando $F(X)$ à cada uma delas.

A $F(X)$ é calculada computacionalmente pela expressão

$$\frac{1}{2}ERF\left(\frac{X}{\sqrt{2}}\right) + \frac{1}{2},$$

onde

$$ERF(X) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

é a função erro (Error Random Function), já incorporada como rotina aos principais programas computacionais.

3) Geração de valores amostrais para os coeficientes com distribuição triangular independentes e dependentes.

Também, nesse caso, os valores t são obtidos a partir da geração de valores com distribuição normal (x), aplicando a função de distribuição normal $F(x)$, obtendo valores (u) de uma variável aleatória com distribuição uniforme. Aplicando a função de distribuição inversa $F_T^{-1}(U) = T$, teremos uma variável com distribuição triangular. A função de distribuição triangular inversa é:

$$F_T^{-1}(U) = \begin{cases} \sqrt{(b-a)(m-a)u} + a, & \text{se } 0 \leq u \leq \frac{m-a}{b-a} \\ b - \sqrt{(u-1)(b-a)(a-b)}, & \text{se } \frac{m-a}{b-a} \leq u \leq 1 \end{cases}$$

Onde a (mínimo), m (moda) e b (máximo) são os parâmetros da variável aleatória com distribuição triangular.

Assim, se U_1 e U_2 são variáveis uniformes dependentes, com correlação ρ_{U_1, U_2} , então $F_T^{-1}(U_1) = T_1$ e $F_T^{-1}(U_2) = T_2$ são variáveis aleatórias triangulares com correlação aproximadamente a mesma $\rho_{T_1, T_2} \cong \rho_{U_1, U_2}$, como pode ser visto pela inspeção dos valores da Tabela 5 e pela seguinte equação:

$$\rho_{T_1, T_2} = -0,0000005369 + 1,011274\rho_{U_1, U_2} \quad -1 \leq \rho_{U_1, U_2} \leq 0$$

(ns) (**) $R^2 = 0,9852$

obtida daqueles dados. Esse procedimento foi realizado em face da dificuldade de se obter essa equação analiticamente. Um asterisco (*) significa que a estimativa do parâmetro da regressão localizado acima é significativamente diferente de 0 pelo teste t-Student, ao nível de 0,05 de probabilidade. Dois asteriscos (**) o é ao nível de 0,01 de probabilidade e um (ns) significa não significativo ao nível 0,05 de probabilidade. O R^2 é o coeficiente de determinação obtido para o ajuste.

- 4) Geração de valores amostrais para os coeficientes com distribuição exponencial independentes e dependentes.

Os valores de E são gerados a partir dos valores gerados com distribuição uniforme (u) e aplicando a função de distribuição exponencial inversa $F_E^{-1}(U) = E$. Assim, se U_1 e U_2 são variáveis uniformes dependentes, com coeficiente de dependência ρ_{U_1, U_2} , então, $F_E^{-1}(U_1) = E_1$ e $F_E^{-1}(U_2) = E_2$, são variáveis aleatórias exponenciais com relação de dependência aproximadamente a mesma $\rho_{E_1, E_2} \cong \rho_{U_1, U_2}$ para valores positivos de ρ_{E_1, E_2} . Equações lineares obtidas por simulação de uma amostra de 2000 valores, para a correlação foram:

$$\begin{array}{lll} \rho_{E_1, E_2} = -0,0345558 + 0,643485\rho_{U_1, U_2} & -1 \leq \rho_{U_1, U_2} \leq 0 \\ (\star) & (\star\star) & R^2 = 0,9925 \end{array}$$

$$\begin{array}{lll} \rho_{E_1, E_2} = -0,016752 + 0,998954\rho_{U_1, U_2} & 0 \leq \rho_{U_1, U_2} \leq 1 \\ (\star) & (\star\star) & R^2 = 0,9956 \end{array}$$

Note que a primeira equação só permite relacionar valores para ρ_{E_1, E_2} , aproximadamente no intervalo $[-0,60;0]$.

A função de distribuição exponencial inversa é:

$$F_E^{-1}(U) = \frac{1}{2} \log(-1/(U - 1))$$

- 5) Geração de valores amostrais para coeficientes com distribuição gama independentes e dependentes.

Nesse caso os valores de G , digamos g , são gerados a partir dos valores gerados com distribuição uniforme (u) e aplicando a seguinte função conforme HAHN & SHAPIRO (1967):

$$g = -\frac{1}{\lambda} \sum_{i=1}^n \ln(1 - u_i), \quad \text{gama}(\eta, \lambda) \quad 0 \leq g < \infty$$

Assim, se U_1 e U_2 são variáveis uniformes dependentes com coeficiente de dependência ρ_{U_1, U_2} , então G_1 e G_2 são variáveis aleatórias gamas com relação de dependência

$$\begin{array}{lll} \rho_{G_1, G_2} = -0,015456 + 0,659961\rho_{U_1, U_2} & -1 \leq \rho_{U_1, U_2} \leq 0 \\ (ns) & (**) & R^2 = 0,9726 \end{array}$$

$$\begin{array}{lll} \rho_{G_1, G_2} = -0,000038177 + 0,989128\rho_{U_1, U_2} & 0 \leq \rho_{U_1, U_2} \leq 1 \\ (ns) & (**) & R^2 = 0,9910 \end{array}$$

Note que a primeira equação só permite relacionar valores para ρ_{G_1, G_2} , aproximadamente no intervalo $[-0,64; 0]$.

6) Geração de valores amostrais para coeficientes com distribuição binomial independentes e dependentes.

Para esse caso os valores de $B(b)$ são gerados a partir dos valores gerados com distribuição uniforme (u) e aplicando a seguinte função:

$$b = \sum_{i=1}^n k_i$$

onde

$$k_i = \begin{cases} 0, & \text{se } u_i > p \\ 1, & \text{se } u_i \leq p, \quad i = 1, 2, \dots, n \end{cases}$$

em que p é a probabilidade de sucesso na binomial (HAHN & SHAPIRO, 1967).

Dessa forma, se U_1 e U_2 são variáveis uniformes dependentes, com coeficiente de dependência ρ_{U_1, U_2} , então B_1 e B_2 são variáveis aleatórias binomiais com relação de dependência:

$$\rho_{B_1, B_2} = -0,005975 + 0,547555\rho_{U_1, U_2} \quad -1 \leq \rho_{U_1, U_2} \leq 1$$

$$(\star\star) \quad (\star) \quad R^2 = 0,9942$$

Note que essa equação só permite relacionar valores para ρ_{B_1, B_2} no intervalo $[-0,5; 0,5]$, aproximadamente.

6.4 Obtenção e Análise dos Resultados

O objetivo do problema de planejamento formulado através da programação linear será maximizar a margem bruta total. Sendo W o conjunto de n coeficientes estocásticos existentes no problema, o processo de simulação terá início com a geração de um primeiro conjunto w de valores amostrais. A solução deste problema, incluindo o conjunto w de valores amostrais, apontará para um determinado plano ótimo, identificado pelos níveis das variáveis x_j do problema. Esse processo será repetido 1000 vezes.

De posse dos conjuntos w de valores amostrais gerados e dos planos ótimos obtidos, o próximo passo será analisar a distribuição da variável margem bruta total.

Se a função objetivo apresentar coeficientes estocásticos incluídos no conjunto w , cada plano ótimo apresentará margem bruta total variável de acordo com as distriuições individuais de cada coeficiente.

A seleção de um único plano ótimo poderá ser orientada através de análise dos parâmetros assimetria, média, moda, variância e probabilidades de ocorrência de valores de margem bruta total superiores a determinados níveis mínimos em cada plano.

6.5 Algoritmo Computacional do Método Proposto

As técnicas de simulação e programação linear, quando voltadas para o estudo de casos reais, somente poderão ser aplicadas se estiverem disponíveis recursos computacionais.

Os métodos de simulação, devido ao grande número de cálculos necessários, exigem a utilização de computadores para incrementar a velocidade de execução e aumentar a exatidão dos resultados. Felizmente, programas específicos para a resolução de problemas de programação linear e linguagens simples, que viabilizam o desenvolvimento de programas de simulação, estão disponíveis na maioria das instituições de ensino e pesquisa.

Nesse trabalho, utilizou-se o procedimento LP (Proc LP)-Linear Programming do módulo OR (Operations Research) do sistema SAS (Statistical Analysis System) para solucionar os modelos de programação linear, bem como o seu módulo básico (BASE) para gerar os valores aleatórios dos coeficientes e o módulo IML (Interactive Matrix Language) para geração de variáveis aleatórias multinormais. Aqui seguiu-se o método das distribuições condicionais.

A listagem completa do programa é apresentada no Apêndice 1.

Na janela OUTPUT do SAS pode-se ver quatro saídas: PROBLEM SUMMARY, SOLUTION SUMMARY, VARIABLE SUMMARY e CONSTRAINT SUMMARY, que contêm informações sobre cada modelo gerado e as respectivas soluções. As soluções são armazenadas em variáveis simples para serem analisadas posteriormente pelo procedimento UNIVARIATE que produz algumas estatísticas básicas.

7 APRESENTAÇÃO DO CASO ESTUDADO

Com o fito de testar a metodologia proposta neste trabalho, tomou-se como estudo do caso o modelo apresentado por SILVA (1995), gentilmente cedido pelo autor.

O modelo de SILVA (1995) apresenta a seguinte formulação matemática:

Maximizar

$$3515x_1 + 3515x_2 + 749x_3 + 3517x_4 + 6832x_5 + 12963x_6 + 7518x_7 + 13251x_8 - 5,50x_9 - x_{18} - x_{20} - x_{21} - x_{22} - 0,08x_{23} = Z \quad (36)$$

sujeito à:

$$x_1 + x_5 + x_6 + x_7 + x_8 \leq 6,5 \quad (37)$$

$$x_2 + x_3 + x_4 + x_6 + x_7 + x_8 \leq 6,5 \quad (38)$$

$$90x_1 + 90x_2 + 64x_3 + 69x_4 + 137x_5 + 255x_6 + 297x_7 + 194x_8 - x_9 \leq 510 \quad (39)$$

$$x_9 \leq 1200 \quad (40)$$

$$-25x_1 - 25x_2 + x_{10} = 0 \quad (41)$$

$$-1800x_3 + x_{11} = 0 \quad (42)$$

$$-25x_4 + x_{12} = 0 \quad (43)$$

$$-45x_5 + x_{13} = 0 \quad (44)$$

$$-38x_6 + x_{14} = 0 \quad (45)$$

$$-16238x_7 + x_{15} = 0 \quad (46)$$

$$-9167x_8 + x_{16} = 0 \quad (47)$$

$$\begin{aligned} -160x_{10} - 0,67x_{11} - 170x_{12} - 180x_{13} - 380x_{14} - 0,56x_{15} \\ -1,6x_{16} + x_{17} \geq 0 \end{aligned} \quad (48)$$

$$\begin{aligned} -46,80x_1 - 46,80x_2 - 46,80x_3 - 46,80x_4 - 46,80x_5 - 93,60x_2 \\ -93,60x_7 - 93,60x_8 + x_{18} \geq 0 \end{aligned} \quad (49)$$

$$\begin{aligned} -5,13x_1 - 5,13x_2 - 4,93x_3 - 4,25x_4 - 3,86x_5 - 15,74x_6 \\ -13,52x_7 - 13,68x_8 + x_{19} \geq 0 \end{aligned} \quad (50)$$

$$-20x_{19} + x_{20} \geq 0 \quad (51)$$

$$-0,001x_{17} + x_{21} \geq 0 \quad (52)$$

$$-0,0067x_{17} + x_{22} \geq 0 \quad (53)$$

$$\begin{aligned} 485x_1 + 485x_2 + 457x_3 + 733x_4 + 1268x_5 + 1476x_6 + 1575x_7 \\ + 1416x_8 + 5,50x_9 + x_{18} + x_{19} - x_{23} - 0,2x_{24} \leq 0 \end{aligned} \quad (54)$$

$$\begin{aligned}
&3515x_1 + 3515x_2 + 749x_3 + 3517x_4 + 6832x_5 + 12963x_6 + 7518x_7 + 13251x_8 \\
&\quad -5,50x_9 - x_{18} - x_{20} - x_{21} - x_{22} - 0,08x_{23} \geq 9538 \quad (55)
\end{aligned}$$

$$\begin{aligned}
&3515x_1 + 3515x_2 + 749x_3 + 3517x_4 + 6832x_5 + 12963x_6 + 7518x_7 + 13251x_8 \\
&\quad -5,50x_9 - x_{18} - x_{20} - x_{21} - x_{22} - 0,08x_{23} - x_{24} = 0 \quad (56)
\end{aligned}$$

$$\begin{aligned}
&485x_1 + 485x_2 + 457x_3 + 733x_4 + 1268x_5 + 1476x_6 + 1575x_7 + 1416x_8 \\
&\quad +5,50x_9 + x_{18} + x_{20} + x_{21} + x_{22} + 0,08x_{23} - x_{25} \geq 1906 \quad (57)
\end{aligned}$$

$$x_6 + x_7 + x_8 \leq 4 \quad (58)$$

$$x_8 \geq 2,1 \quad (59)$$

$$x_4 \leq 1,5 \quad (60)$$

$$x_3 \geq 0,5 \quad (61)$$

$$x_6 \geq 1,0 \quad (62)$$

A Tabela 6 apresenta a codificação e definição das atividades utilizadas no modelo de programação linear.

A Tabela 7 apresenta a definição da função objetivo e das restrições utilizadas no modelo de programação linear.

Tabela 6. Codificação e definição das atividades utilizadas no modelo de programação linear.

| Código | Definição | Unidade |
|----------|--|-------------|
| x_1 | área cultivada c/ abóbora no 1º semestre | hectare |
| x_2 | área cultivada c/ abóbora no 2º semestre | hectare |
| x_3 | área cultivada c/ feijão no 2º semestre | hectare |
| x_4 | área cultivada c/ melancia no 2º semestre | hectare |
| x_5 | área cultivada c/ tomate no 1º semestre | hectare |
| x_6 | área cultivada c/ banana | hectare |
| x_7 | área cultivada c/ goiaba | hectare |
| x_8 | área cultivada c/ manga | hectare |
| x_9 | mão-de-obra contratada | dia/homem |
| x_{10} | produção de abóbora | tonelada |
| x_{11} | produção de feijão | quilograma |
| x_{12} | produção de melancia | tonelada |
| x_{13} | produção de tomate | tonelada |
| x_{14} | produção de banana | tonelada |
| x_{15} | produção de goiaba | quilograma |
| x_{16} | produção de manga | quilograma |
| x_{17} | renda bruta | reais (R\$) |
| x_{18} | parcela de tarifa d'água relativa à amortização dos investimentos públicos | reais (R\$) |
| x_{19} | consumo anual de água | $1000m^3$ |
| x_{20} | valor do consumo de água | reais (R\$) |
| x_{21} | despesas com FUNRURAL | reais (R\$) |
| x_{22} | custo com comercialização | reais (R\$) |
| x_{23} | demanda por crédito bancário | reais (R\$) |
| x_{24} | margem bruta | reais (R\$) |
| x_{25} | despesas | reais (R\$) |

Tabela 7. Definição da função objetivo e das restrições utilizadas no modelo de programação linear.

| Descrição | Unidade |
|--|---|
| 1 - Maximização da margem bruta (Z) | reais (R\$) |
| 2 - Utilização da terra no 1º semestre | hectare (ha) |
| 3 - Utilização da terra no 2º semestre | hectare (ha) |
| 4 - Utilização de mão-de-obra | dias/homem (D/H) |
| 5 - Disponibilidade de mão-de-obra | dias/homem (D/H) |
| 6 - Produção de abóbora | tonelada (t) |
| 7 - Produção de feijão | quilograma (kg) |
| 8 - Produção de melancia | tonelada (t) |
| 9 - Produção de tomate | tonelada (t) |
| 10 - Produção de banana | tonelada (t) |
| 11 - Produção de goiaba | quilograma (kg) |
| 12 - Produção de manga | quilograma (kg) |
| 13 - Valor da produção total | reais (R\$) |
| 14 - Amortização total | reais (R\$) |
| 15 - Consumo de água | mil metros cúbicos (1000 m ³) |
| 16 - Valor do consumo de água | reais (R\$) |
| 17 - Despesas com FUNRURAL | reais (R\$) |
| 18 - Custo com comercialização | reais (R\$) |
| 19 - Crédito | reais (R\$) |
| 20 - Renda mínima | reais (R\$) |
| 21 - Margem bruta | reais (R\$) |
| 22 - Despesas | reais (R\$) |
| 23 - Área máxima com culturas perenes | hectare (ha) |
| 24 - Área mínima com manga | hectare (ha) |
| 25 - Área mínima com melancia no 2º semestre | hectare (ha) |
| 26 - Área mínima com feijão no 2º semestre | hectare (ha) |
| 27 - Área mínima com banana | hectare (ha) |

O estudo foi realizado no Perímetro de Irrigação Senador Nilo Coelho que está localizado à margem esquerda do rio São Francisco, nos municípios de Casa Nova (BA) e Petrolina (PE), em plena região semi-árida, onde a precipitação média anual é da ordem de 613,8 mm, uma das regiões mais secas do nordeste brasileiro. Trata-se do maior projeto de irrigação sob a administração da Companhia de Desenvolvimento do Vale do São Francisco (CODEVASF), com área irrigável de 20.018 ha. Conta, atualmente, com 1.432 famílias de pequenos produtores (colonos) assentados, com parcelas médias de 6,5 ha irrigáveis e 1,0 ha de sequeiro, representando cerca de 70% da área irrigada já implantada. Os 30% restantes de áreas já irrigadas estão distribuídos entre 131 pequenas, médias e grandes empresas, com glebas de 12 a 490 hectares. O tipo de irrigação é por aspersão (SILVA, 1995 e BISERRA et al., 1995).

Os autores identificaram três modelos atuais de exploração, partindo-se da diferenciação das unidades produtivas para o posterior agrupamento dessas em unidades homogêneas, considerando-se os critérios de área total e de área efetivamente cultivada em cada unidade, bem como a combinação de culturas exploradas atualmente e suas respectivas áreas.

Consideraram, também, três diferentes cenários em relação ao custo da água e níveis mínimos de renda para o colono e sua família.

O caso tomado aqui foi o Modelo de Exploração 1 que representa as unidades de produção que cultivam abóbora, feijão, melancia, tomate e manga. Em termos de participação relativa, as culturas deste modelo ocupam, em 1995, cerca de 51,3% da área total cultivada no perímetro SILVA (1995). E o cenário 1 foi tomado aqui por refletir a situação atual, ou seja, a cobrança de uma tarifa d'água altamente subsidiada.

Consideraram-se aleatórios alguns coeficientes da função objetivo e da restrição consumo de água ($1000 m^3$), como mostrado na Tabela 8. Os valores médios dos parâmetros foram obtidos do trabalho de SILVA (1995) e as variâncias, bem como as correlações, foram estabelecidas com base no estudo de séries históricas

pesquisadas em outros trabalhos como BIZARRIA (1993) e SILVA (1988), e as distribuições foram obtidas em estudos feitos por BISERRA et al. (1995) e em comunicação pessoal com o professor José Valdeci Biserra da Universidade Federal do Ceará.

Foram adotados os seguintes valores para coeficientes de correlação não nulos entre coeficientes aleatórios distintos:

$$\rho_{c_1,c_2} = 0,5; \rho_{c_1,c_3} = 0,375; \rho_{c_1,c_4} = 0,5; \rho_{c_1,c_5} = 0,4375; \rho_{c_1,c_6} = 0,4375$$

$$\rho_{c_1,c_7} = 0,375; \rho_{c_1,c_8} = 0,4375; \rho_{c_1,a_{14,1}} = 0,5625;$$

$$\rho_{c_2,c_3} = 0,375; \rho_{c_2,c_4} = 0,5; \rho_{c_2,c_5} = 0,3125; \rho_{c_2,c_6} = 0,4375;$$

$$\rho_{c_2,c_7} = 0,375; \rho_{c_2,c_8} = 0,4375; \rho_{c_2,a_{14,2}} = 0,5625;$$

$$\rho_{c_3,c_4} = 0,3125; \rho_{c_3,c_5} = 0,375; \rho_{c_3,c_6} = 0,25; \rho_{c_3,c_7} = 0,4375;$$

$$\rho_{c_3,c_8} = 0,3125; \rho_{c_3,a_{14,3}} = 0,5;$$

$$\rho_{c_4,c_5} = 0,5; \rho_{c_4,c_6} = 0,5; \rho_{c_4,c_7} = 0,4375; \rho_{c_4,c_8} = 0,4375;$$

$$\rho_{c_4,a_{14,4}} = 0,5625;$$

$$\rho_{c_5,c_6} = 0,375; \rho_{c_5,c_7} = 0,3125; \rho_{c_5,c_8} = 0,3125; \rho_{c_5,a_{14,5}} = 0,59375;$$

$$\rho_{c_6,c_7} = 0,4375; \rho_{c_6,c_8} = 0,5; \rho_{c_6,a_{14,6}} = 0,5;$$

$$\rho_{c_7,c_8} = 0,4375; \rho_{c_7,a_{14,7}} = 0,4375;$$

$$\rho_{c_8,a_{14,8}} = 0,4375;$$

O valor $\rho_{c_1,c_2} = 0,5$ significa que a correlação, entre as variáveis aleatórias com distribuição normal X_1 (margem bruta da abóbora no primeiro semestre) e X_2 (margem bruta da abóbora no segundo semestre), é positiva e igual a 0,5.

Tabela 8. Distribuição de probabilidade dos coeficientes aleatórios assumidos no modelo de programação linear

| Coeficiente | Atividade | Localização | Distribuição | Parâmetros¹ |
|--------------------|------------------|--------------------|---------------------|-------------------------------|
| c_1 | x_1 | Função objetivo | Normal | (3515;161,7) |
| c_2 | x_2 | Função objetivo | Normal | (3515;161,7) |
| c_3 | x_3 | Função objetivo | Triangular | (666,6;749;872,6) |
| c_4 | x_4 | Função objetivo | Triangular | (2673;3517;4080) |
| c_5 | x_5 | Função objetivo | Triangular | (4099;6832;9565) |
| c_6 | x_6 | Função objetivo | Triangular | (8100;12963;17825) |
| c_7 | x_7 | Função objetivo | Triangular | (5262,6;7518;9773,4) |
| c_8 | x_8 | Função objetivo | Triangular | (9607;13251;16895) |
| $a_{14,1}$ | x_1 | Consumo de água | Uniforme | (3,26;5,13;7) |
| $a_{14,2}$ | x_2 | Consumo de água | Uniforme | (3,26;5,13;7) |
| $a_{14,3}$ | x_3 | Consumo de água | Uniforme | (3,93;4,93;5,93) |
| $a_{14,4}$ | x_4 | Consumo de água | Uniforme | (2,25;4,25;6,25) |
| $a_{14,5}$ | x_5 | Consumo de água | Uniforme | (3,36;3,86;4,36) |
| $a_{14,6}$ | x_6 | Consumo de água | Uniforme | (14,74;15,74;16,74) |
| $a_{14,7}$ | x_7 | Consumo de água | Uniforme | (12,12;13,52;14,92) |
| $a_{14,8}$ | x_8 | Consumo de água | Uniforme | (13,28;13,68;14,08) |

¹ Para a distribuição Normal os parâmetros são (média; variância).

Para a distribuição Triangular os parâmetros são (mínimo; moda; máximo).

Para a distribuição Uniforme os parâmetros são (mínimo; média; máximo).

Tabela 9. Área cultivada (ha) com as atividades que compõem os planos atual, proposto pelo modelo determinístico e obtido pela simulação com amostra de tamanho mil.

| Atividade | Plano atual | Plano proposto determinista | Simulação | | | | | | | |
|-----------|-------------|-----------------------------|-----------------|--------|------|------|-----------------|-------|------|------|
| | | | Sem dependência | | | | Com dependência | | | |
| | | | Min. | Médio | Max. | Moda | Min. | Médio | Max. | Moda |
| x_1 | 1,5 | 0,0 | 0,0 | 0,0075 | 2,5 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| x_2 | 1,0 | 0,5 | 0,5 | 1,2 | 2,9 | 0,5 | 0,5 | 1,19 | 2,9 | 0,5 |
| x_3 | 1,5 | 0,5 | 0,5 | 0,5 | 0,5 | 0,5 | 0,5 | 0,5 | 0,5 | 0,5 |
| x_4 | 1,0 | 1,5 | 0,0 | 0,0078 | 1,5 | 1,5 | 0,0 | 0,81 | 1,5 | 1,5 |
| x_5 | 2,0 | 2,5 | 0,0 | 2,5 | 3,4 | 2,5 | 2,5 | 2,51 | 3,4 | 2,5 |
| x_6 | - | 1,0 | 1,0 | 1,3 | 1,9 | 1,0 | 1,0 | 1,32 | 1,9 | 1,0 |
| x_7 | - | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| x_8 | 3,0 | 3,0 | 2,1 | 2,6 | 3,0 | 3,0 | 2,1 | 2,67 | 3,0 | 3,0 |

O valor $\rho_{c_1,c_3} = 0,375$ significa que a correlação, entre as variáveis aleatórias X_1 com distribuição normal e X_3 (margem bruta do feijão no segundo semestre) com distribuição triangular, é positiva igual a 0,375.

O valor $\rho_{c_1,a_{14,1}} = 0,5625$ significa que a correlação, entre as variáveis aleatórias X_1 com distribuição normal e X_9 (consumo anual de água pela abóbora no primeiro semestre) com distribuição uniforme, é positiva igual a 0,5625.

Por fim o valor $\rho_{c_3,a_{14,3}} = 0,5$ significa que a correlação, entre as variáveis aleatórias X_3 com distribuição triangular e X_{11} (consumo anual de água pelo feijão no segundo semestre) com distribuição uniforme, é positiva igual a 0,5.

A validação do método de simulação se baseou na análise de alguns parâmetros estatísticos dos valores amostrais gerados para os coeficientes aleatórios. Foram gerados mil valores amostrais para cada coeficiente aleatório.

A Tabela 9 mostra o plano atual, o plano proposto com base no modelo determinístico e os resultados da simulação, para as atividades “culturas anuais e permanente”.

Nota-se, pelos resultados da simulação, tomando a moda, que os planos propostos são praticamente os mesmos daqueles obtidos pela solução do

modelo determinístico, estando de acordo com os resultados encontrados por RODRIGUES (1987), embora sejam diferentes do plano atual.

A variância da variável margem bruta (Z) é menor no caso da simulação com coeficientes independentes do que quando eles foram considerados dependentes.

As estatísticas dos resultados da função objetivo são apresentadas na Tabela 10.

As estatísticas para a variável margem bruta (Z) mostram que essa variável apresenta uma certa assimetria negativa confirmada pelo teste de normalidade devido a Shapiro-Wilk, que mostra que a distribuição é estatisticamente diferente da normal ($\text{Pr} < W = 0,0159$) para o caso da simulação com dependência e normalidade dos dados ($\text{Pr} < W = 0,4058$) para o caso da simulação sem dependência (Figuras 6 e 7).

A matriz de correlações deve ser semidefinida positiva. Uma idéia do quanto os elementos dessa matriz estão associados pode ser exemplificada pelo fato de que se tomarmos três variáveis X_1, X_2 e X_3 e se X_1 e X_2 estão altamente correlacionados e X_2 também o está com X_3 , X_1 não pode ter baixa correlação com X_3 . Há neste caso, menor liberdade de atribuímos valor à esta última correlação.

Os resultados obtidos para as variáveis x_j em cada simulação estabeleceriam um plano ótimo caso prevalecessem efetivamente os valores amostrais gerados para os coeficientes nessa simulação.

Denominamos estado da natureza ao conjunto de valores amostrais obtidos x_{ij} em cada simulação. Verificou-se que o método proposto resulta em um número menor de planos ótimos quando se considera a estrutura de dependência.

O programa utilizado, escrito em SAS, encontra-se no Apêndice 1 e uma amostra dos resultados das simulações do problema, está apresentada no Apêndice 2.

O problema apresentado e submetido ao processo de simulação, apresenta seis e sete planos diferentes de se enfrentar a aleatoriedade das margens

Tabela 10. Estatísticas simples para a margem bruta obtidas dos resultados de simulação com e sem dependência entre os coeficientes do modelo de programação linear.

| Estatísticas | Sem dependência | Com dependência |
|--------------------------|------------------------------------|------------------------------------|
| Média | 70.246,66 | 69.953,42 |
| Moda | 70.100,00 | 69.650,00 |
| Mediana | 70.263,30 | 69.919,51 |
| Assimetria | -0,02 | - 0,01 |
| Curtose | -0,18 | -0,18 |
| Desvio padrão | 5.681,98 | 7.609,66 |
| Variância | 32.284.901,00 | 57.906.865,00 |
| Erro padrão da média | 179,68 | 240,64 |
| Teste de normalidade | $P_r < W : 0,4058$ $W = 0,9861$ | $P_r < W : 0,0159$ $W = 0,9819$ |
| Coefficiente de variação | 8,08% | 10,87% |
| Primeiro quartil | 66.470,87 | 64.883,31 |
| Terceiro quartil | 74.068,13 | 75.021,61 |
| Mínimo | 53.203,90 | 49.196,97 |
| Máximo | 88.431,76 | 90.944,62 |

Tabela 11. Frequência dos planos propostos obtidos do modelo de programação linear através da simulação com dependência entre os coeficientes aleatórios.

| Plano | X_1 | X_2 | X_3 | X_4 | X_5 | X_6 | X_7 | X_8 | Frequência (%) | Margem bruta média | Desvio padrão |
|--------------|-------|-------|-------|-------|-------|-------|-------|-------|-----------------------|---------------------------|----------------------|
| 1 | 0,0 | 0,5 | 0,5 | 1,5 | 2,5 | 1,0 | 0,0 | 3,0 | 33,8 | 68.409,37 | 7.520,62 |
| 2 | 0,0 | 2,0 | 0,5 | 0,0 | 2,5 | 1,0 | 0,0 | 3,0 | 29,6 | 69.609,48 | 7.707,38 |
| 3 | 0,0 | 0,5 | 0,5 | 1,5 | 2,5 | 1,9 | 0,0 | 2,1 | 19,6 | 71.180,09 | 6.987,55 |
| 4 | 0,0 | 2,0 | 0,5 | 0,0 | 2,5 | 1,9 | 0,0 | 2,1 | 15,9 | 72.564,78 | 7.666,80 |
| 5 | 0,0 | 1,4 | 0,5 | 1,5 | 3,4 | 1,0 | 0,0 | 2,1 | 00,6 | 65.752,72 | 4.375,88 |
| 6 | 0,0 | 2,9 | 0,5 | 0,0 | 3,4 | 1,0 | 0,0 | 2,1 | 00,5 | 68.606,17 | 2.512,98 |

brutas para os casos com (tabela 11) e sem (tabela 12) dependência entre os coeficientes aleatórios do modelo de programação linear.

O plano proposto determinístico apresenta como resultado o plano 1, indicado nas tabelas 11 e 12. Este plano maximiza a margem bruta em 33,8% e 32,0% dos prováveis estados da natureza para os casos simulados com e sem dependência, ocorrendo neste caso, uma predominância.

Dependendo da aversão do colono à variação da margem bruta, o plano 1 pode ser mais atraente, devido à sua alta probabilidade de se manter como plano ótimo.

Tabela 12. Frequência dos planos propostos obtidos do modelo de programação linear através da simulação sem dependência entre os coeficientes aleatórios.

| Plano | X_1 | X_2 | X_3 | X_4 | X_5 | X_6 | X_7 | X_8 | Frequência (%) | Margem bruta média | Desvio padrão |
|--------------|-------|-------|-------|-------|-------|-------|-------|-------|-----------------------|---------------------------|----------------------|
| 1 | 0,0 | 0,5 | 0,5 | 1,5 | 2,5 | 1,0 | 0,0 | 3,0 | 32,0 | 70.159,60 | 5.701,06 |
| 2 | 0,0 | 2,0 | 0,5 | 0,0 | 2,5 | 1,0 | 0,0 | 3,0 | 27,6 | 70.322,64 | 6.129,06 |
| 3 | 0,0 | 0,5 | 0,5 | 1,5 | 2,5 | 1,9 | 0,0 | 2,1 | 19,6 | 70.429,8 | 5.126,59 |
| 4 | 0,0 | 2,0 | 0,5 | 0,0 | 2,5 | 1,9 | 0,0 | 2,1 | 18,3 | 70.683,61 | 5.484,88 |
| 5 | 0,0 | 2,9 | 0,5 | 0,0 | 3,4 | 1,0 | 0,0 | 2,1 | 01,5 | 66.468,97 | 4.309,33 |
| 6 | 0,0 | 1,4 | 0,5 | 1,5 | 3,4 | 1,0 | 0,0 | 2,1 | 00,7 | 65.713,13 | 2.799,29 |
| 7 | 2,5 | 0,5 | 0,5 | 1,5 | 0,0 | 1,9 | 0,0 | 2,1 | 00,3 | 63.388,68 | 7.146,22 |

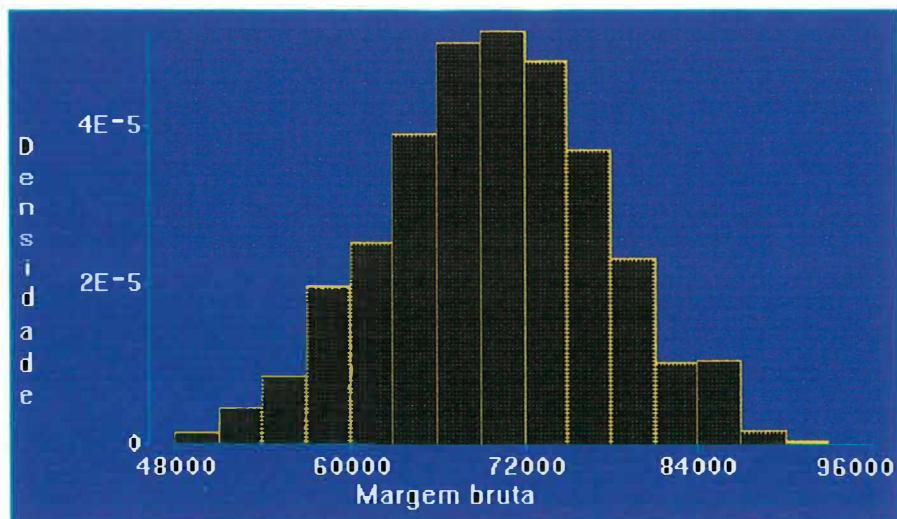


Figura 6. Distribuição de frequência da margem bruta relativa ao modelo de programação linear simulado com dependência entre os coeficientes.

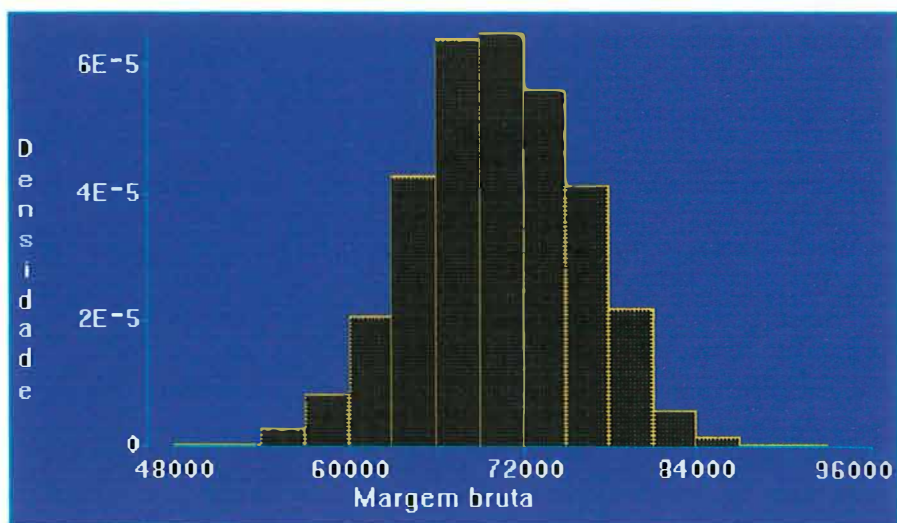


Figura 7. Distribuição de frequência da margem bruta relativa ao modelo de programação linear simulado sem dependência entre os coeficientes.

8 CONCLUSÕES E SUGESTÕES

- Pode-se gerar amostras com estrutura de dependência para distribuições como: uniforme, triangular, exponencial, gama, poisson etc.
- A inclusão de risco em modelos de planejamento agrícola através de técnicas de programação linear o torna mais realistas, embora neste caso particular os resultados não tenham mostrado esta melhoria.
- O uso de simulação mostrou-se bastante útil como instrumento de análise dos efeitos da aleatoriedade dos coeficientes estocásticos sobre o resultado do processo de otimização implementada pela programação linear.
- O fato de se ter obtido um conjunto pequeno de soluções para um número relativamente grande de simulações pode não se repetir quando da análise de outros problemas.
- O uso da simulação sem dependência entre os coeficientes aleatórios subestimou o risco quando comparado com a simulação com dependência, para esse caso.
- A maior dificuldade encontrada neste trabalho, reside na construção de uma matriz de correlações (dependência) que seja semidefinida positiva, exigência esta da metodologia empregada para simular as variáveis. O uso de séries históricas para geração da matriz de correlações garante que a condição seja satisfeita.
- Sugere-se que outros problemas de programação linear não determinísticos sejam estudados através do processo de simulação em que outras origens da distribuição uniforme multivariada sejam utilizadas visando obter distribuições conjuntas melhor comportadas. Aqui usou-se como origem da distribuição conjunta uniforme

a distribuição normal. Melhor comportada no sentido da distribuição condicional $(U_1|U_2)$ ter um comportamento mais uniforme.

- Sugere-se que se faça um estudo analítico para as relações de dependência entre os coeficientes de correlação que envolvem as distribuições exponencial, triangular, gama e binomial e a distribuição uniforme.

REFERÊNCIAS BIBLIOGRÁFICAS

- AGRAWAL, R.C.; HEADY, E.O. **Operations research methods for agricultural decisions** Ames: The Iowa State University Press, 1972. 303p.
- ANDERSON, JR.; DILLON, J.L.; HARDAKER, J.B. **Agricultural decisions analysis**: Ames: The Iowa State University Press, 1977. 344p.
- BARNETT, V. Some bivariate uniform distributions. **Communications in Statistics - Theory and Methods**. v.9, n.4, p.453-61, 1980.
- BISERRA, J.V.; ABDIAS FILHO, N.; CÂMARA, S.F. ET AL. Rentabilidade da irrigação Pública no Nordeste, sob Condições de risco. **Revista Econômica do Nordeste**, v.26, n.2, abr./jun. p.239-63, 1995.
- BIZARRIA, S.A.M. Avaliação da seleção de sistemas de produção agrícola sob condição de risco. Fortaleza, 1993. 149p. Dissertação (Mestrado) - Universidade Federal do Ceará.
- BOSWELL, S.D.; GORE, G.P.; PATIL, G.P.; TAILLIE, C. The art of computer generation of random variables. In: RAO, C.R., ed. **Computational Statistics**. Amsterdam: North-Holland, 1993. p.661-721. (Handbook of Statistics, 9).
- BOUSSARD, J.M.; PETIT, M. Representation of farmers behavior under uncertainty, with a focus-loss constraint. **Journal Farm Economics**, v.49, n.4, p.869-80, Nov. 1967.

- CALATRAVA, J.; DOMINGO, J. Consideraciones sobre el empleo de la simulación a la resolución de programas estocásticos: planificación de cultivos de una explotación hortícola de la costa mediterránea de Andalucía. **Anales del INIA**, v.6, p.35-54, 1982.
- COCKS, K.D. Discrete stochastic programming. **Management Science**, v.15, n.1, p.72-9, Sept. 1968.
- GODOI, C.R.M. Um algoritmo eficiente para simulação de vetores com distribuição multinormal. **Ciência e Cultura**, v.30, n.6, p.701-5, 1978.
- HAN, G.J.; SHAPIRO, S.S. **Statistical models in engineering**. New York: John Wiley, 1967. 355p.
- HADLEY, G. **Nonlinear and dynamic programming**. Reading; Addison-Wesley, 1964. 484p.
- HAZZEL, P.B.R. Game theory: on extension of its application to farm planning under uncertainty. **Journal of Agricultural Economics**, v.21, n.2, p.239-52, May, 1970.
- HAZZEL, P.B.R. A Linear alternative to quadratic and semi variance programming for farm planning under uncertainty. **Journal of Agricultural Economics**, v.53, n.1, p.53-62, 1971.
- HILLIER, S.; LIEBERMAN, G.J. **Introdução à pesquisa operacional**, Rio de Janeiro: Ed. Campus, 1988. 805p.
- JOHNSON, M.E. **Multivariate statistical simulation**, New York: John Wiley, 1987. 230p.
- KLEIJNEN, J.P.C. **Statistical techniques in simulation**, New York: Marcel Dekker, 1974. pt.1, 285p.

- LANZER, E.A. **Programação linear: Conceitos e aplicações.** Rio de Janeiro: IPEA, INPES, 1982. 258p.
- LEONTIFF, W. **The structure of the American Economy, 1919-1939.** 2.ed. Fair Lawun: Oxford University Press, 1951. p.38-45.
- MARKOWITZ, H.M. Portfolio selection. **Journal of Finance**, v.7, p.77-91, May, 1952.
- MARKOWITZ, H.M. **Portfolio selection: efficient diversification of investments.** New York: John Willey, 1959. p.84-96.
- MCINERNEY, J.P. Linear programing and game theory models - some extensions. **Journal of Agricultural Economics**, v.20, p.269-78, May, 1969.
- NEUMANN, J. VON; MORGENSTERN, O. **Theory of games and economic behavior.** 3.ed. Princeton: Princenton University Press. 1953.
- RAE, A.N. An empirical application and evaluation of discrete stochastic programming in farm management. **Journal of Agricultural Economics**, v.53, n.4, p.625-37, 1971a.
- RAE, A.N. Stochastic programming, utility, and sequential decision problems in farm management. **Journal of Agricultural Economics**, v.53, n.3, p.448-60, 1971b.
- RAMSEY, F.P. **Truth and probability, in the logical foundations of mathematics and other essays,** Kegan Paul. Reprinted by Kyburg and smokier, 1964, p.61-92.

- RODRIGUES, L.C.E. Planejamento agropecuário através de um modelo de Programação linear não determinista. Piracicaba, 1987. 83p. Dissertação (Mestrado) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo.
- SAVAGE, L.J. **The foundations of statistics**. New York: Dover, 1954.
- SILVA, J.H.F. Rentabilidade de uma Proposta de Produção para as Unidades de Exploração do Perímetro de Irrigação Senador Nilo Coelho (BA/PE). Fortaleza, 1995. 98p. Dissertação (Mestrado) - Universidade Federal do Ceará.
- SILVA, J.R. Planejamento Agrícola sob Condições de risco para pequenas propriedades da Zona semi-Árida do Estado do Ceará. Fortaleza, 1988. 160p. Dissertação (Mestrado) - Universidade Federal do Ceará.
- SILVA, L.P.A. Fundamentos da teoria da decisão. In: SIMPÓSIO BRASILEIRO DE PROBABILIDADE E ESTATÍSTICA, 9., São Paulo, 1990. **Anais**. São Paulo: USP, IME. 81p.
- SOARES, F.B. Programação da Empresa em Condições de Incerteza: Uma possível Abordagem. **Economia**, v.1, n.2, p.319-35, maio, 1977.
- WEBER, J.E. **Matemática para economia e administração**. 3.ed. São Paulo: Harbra, 1977. 226p.
- WILKS, S.S. **Mathematical statistics**. New York: John Wiley, 1962. 644p.

APÊNDICE 1

**PROGRAMA DESENVOLVIDO EM LINGUAGEM SAS PARA SIMULAR
COEFICIENTES ALEATÓRIOS E SOLUCIONAR O PROBLEMA DE
PROGRAMAÇÃO LINEAR**

```

proc iml;
wrksize=100000;
K=16;      /*K: Número de coeficientes aleatórios no modelo de PL */
N=1000;    /*N: Tamanho da amostra à ser simulada */
M={0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0}; /*Vetor de médias */
      /*S: Matriz de correlações */
S={
  1    0.5 0.375 0.5 0.4375 0.4375 0.375
  0.4375 0.5625 0 0 0 0 0
  0 0,
  0.5 1 0.375 0.5 0.3125 0.4375 0.375
  0.4375 0 0.5625 0 0 0 0
  0 0,
  0.375 0.375 1 0.3125 0.375 0.25 0.4375
  0.3125 0 0 0.5 0 0 0
  0 0,
  0.5 0.5 0.3125 1 0.5 0.5 0.4375
  0.4375 0 0 0 0.5625 0 0
  0 0,
  0.4375 0.3125 0.375 0.5 1 0.375 0.3125
  0.3125 0 0 0 0 0.59375 0
  0 0,
  0.4375 0.4375 0.25 0.5 0.375 1 0.4375
  0.5 0 0 0 0 0 0.5
  0 0,
  0.375 0.375 0.4375 0.4375 0.3125 0.4375 1
  0.4375 0 0 0 0 0 0
  0.4375 0,
  0.4375 0.4375 0.3125 0.4375 0.3125 0.5 0.4375
  1 0 0 0 0 0 0
  0 0.4375,
  0.5625 0 0 0 0 0 0
  0 1 0 0 0 0 0
  0 0,
  0 0.5625 0 0 0 0 0
  0 0 1 0 0 0 0
  0 0,
  0 0 0.5 0 0 0 0
  0 0 0 1 0 0 0
  0 0,
  0 0 0 0.5625 0 0 0
  0 0 0 0 1 0 0
  0 0,
  0 0 0 0 0.59375 0 0

```

```

0    0    0    0    0    1    0
0    0,
0    0    0    0    0    0.5    0
0    0    0    0    0    0    1
0    0,
0    0    0    0    0    0    0.4375
0    0    0    0    0    0    0
1    0,
0    0    0    0    0    0    0
0.4375  0    0    0    0    0    0
0    1};

```

```
/* Procedimento para simular normal multidimensional */
```

```

X=shape(0,K,N);
ME=0; SI=1;
DO I=1 TO K;
  DO J=1 TO N;
    if I>1
      then
        do;
          ME=M[I]+(S[1:I-1,I])`*(inv(S[1:I-1,1:I-1])*(X[1:I-1,J]-M[1:I-1]));
          SI=S[I,I]-(S[1:I-1,I])`*(inv(S[1:I-1,1:I-1])*(S[1:I-1,I]));
        end;
        X[I,J]=ME+NORMAL(0)*SQRT(SI);
      end;
  end;
END;

END;
Z=X;

```

```

varnames='X1':'X16';
/* Cria arquivo SASdataset (NOVO) à partir do módulo IML */

```

```

create NOVO from Z[colname=varnames];
append from Z;
quit; /* Sai do módulo IML */

```

```

/* Simulação dos coeficientes aleatórios do modelo de PL */
data PERIME; set NOVO;
options ps=66 ls=75;

```

```

/* Geração dos coeficientes aleatórios com distribuição normal */
X1=3515+402.125*X1;
X2=3515+402.125*X2;

```

```
/* Definição dos parâmetros dos coeficientes aleatórios com distribuição
```

```

triangular */
a1=666.6; m1=749; b1=872.6;
a2=2673; m2=3517; b2=4080;
a3=4099; m3=6832; b3=9565;
a4=8100; m4=12963; b4=17825.5;
a5=5262.6; m5=7518; b5=9773.4;
a6=9607; m6=13251; b6=16895;

/* Geração inicial das variáveis uniformes */
y1=erf(sqrt(2)*x3/2)/2+0.5;
y2=erf(sqrt(2)*x4/2)/2+0.5;
y3=erf(sqrt(2)*x5/2)/2+0.5;
y4=erf(sqrt(2)*x6/2)/2+0.5;
y5=erf(sqrt(2)*x7/2)/2+0.5;
y6=erf(sqrt(2)*x8/2)/2+0.5;
y7=erf(sqrt(2)*x9/2)/2+0.5;
y8=erf(sqrt(2)*x10/2)/2+0.5;
y9=erf(sqrt(2)*x11/2)/2+0.5;
y10=erf(sqrt(2)*x12/2)/2+0.5;
y11=erf(sqrt(2)*x13/2)/2+0.5;
y12=erf(sqrt(2)*x14/2)/2+0.5;
y13=erf(sqrt(2)*x15/2)/2+0.5;
y14=erf(sqrt(2)*x16/2)/2+0.5;

/* Geração dos coeficientes aleatórios com distribuição triangular */
if (y1>=0) and (y1<=(m1-a1)/(b1-a1))
  then t1=sqrt((b1-a1)*(m1-a1)*y1)+a1;
if (y1>(m1-a1)/(b1-a1)) and (y1<=1)
  then t1= b1-sqrt((y1-1)*(b1-m1)*(a1-b1));

if (y2>=0) and (y2<=(m2-a2)/(b2-a2))
  then t2=sqrt((b2-a2)*(m2-a2)*y2)+a2;
if (y2>(m2-a2)/(b2-a2)) and (y2<=1)
  then t2= b2-sqrt((y2-1)*(b2-m2)*(a2-b2));

if (y3>=0) and (y3<=(m3-a3)/(b3-a3))
  then t3=sqrt((b3-a3)*(m3-a3)*y3)+a3;
if (y3>(m3-a3)/(b3-a3)) and (y3<=1)
  then t3= b3-sqrt((y3-1)*(b3-m3)*(a3-b3));

if (y4>=0) and (y4<=(m4-a4)/(b4-a4))
  then t4=sqrt((b4-a4)*(m4-a4)*y4)+a4;
if (y4>(m4-a4)/(b4-a4)) and (y4<=1)
  then t4= b4-sqrt((y4-1)*(b4-m4)*(a4-b4));

if (y5>=0) and (y5<=(m5-a5)/(b5-a5))
  then t5=sqrt((b5-a5)*(m5-a5)*y5)+a5;

```



```

if (y5>(m5-a5)/(b5-a5)) and (y5<=1)
  then t5= b5-sqrt((y5-1)*(b5-m5)*(a5-b5));

if (y6>=0) and (y6<=(m6-a6)/(b6-a6))
  then t6=sqrt((b6-a6)*(m6-a6)*y6)+a6;
if (y6>(m6-a6)/(b6-a6)) and (y6<=1)
  then t6= b6-sqrt((y6-1)*(b6-m6)*(a6-b6));

/* Geração dos coeficientes aleatórios com distribuição uniforme */
U1=-(3.26+Y7*(7-3.26));
U2=-(3.26+Y8*(7-3.26));
U3=-(3.93+Y9*(5.93-3.93));
U4=-(2.25+Y10*(6.25-2.25));
U5=-(3.36+Y11*(4.36-3.36));
U6=-(14.74+Y12*(16.74-14.74));
U7=-(12.12+Y13*(14.92-12.12));
U8=-(13.28+Y14*(14.08-13.28));

/* Cálculo das correlações entre os coeficientes aleatórios gerados */
proc corr data=PERIME;
var x1 x2 t1 t2 t3 t4 t5 t6 u1 u2 u3 u4 u5 u6 u7 u8;
run;

/* Cálculo de algumas estatísticas simples para os coeficientes aleatórios */
proc univariate plot data=PERIME;
var x1 x2 t1 t2 t3 t4 t5 t6 u1 u2 u3 u4 u5 u6 u7 u8;
run;

/*-----*/
/*SOLUCAO DO PROBLEMA DE PROGRAMACAO LINEAR      */
/*-----*/
data IRRIGA;
input _id_ $
      AB1 AB2 FE2 ME2 TO1 BA GO MG MOC PAB PFE PME PTO PBA PGO PMG
      RB AMORT AGUA VAGUA FUR COM CRED MB DESP
      _type_ $ _rhs_ @@;
cards;
OBJET  3515 3515 749 3517 6832 12963 7518 13251 -5.50 0 0 0 0 0 0 0
      0 -1 0 -1 -1 -1 -0.08 0 0 MAX .
TERRA1  1 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0
      0 0 0 0 0 0 0 0 0 0 LE 6.5
TERRA2  0 1 1 1 0 1 1 1 0 0 0 0 0 0 0
      0 0 0 0 0 0 0 0 0 0 LE 6.5
UMAOB   90 90 64 69 137 255 297 194 -10 0 0 0 0 0 0
      0 0 0 0 0 0 0 0 0 0 LE 510
DMAOB   0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
      0 0 0 0 0 0 0 0 0 0 LE 1200

```

```

PROAB  -25 -25 0 0 0 0 0 0 0 0 01000000
        0 0 0 0 0 0 0 0 0 0LE 0
PROFE  0 0 -1800 0 0 0 0 0 0 0 00100000
        0 0 0 0 0 0 0 0 0 0LE 0
PROME  0 0 0 -25 0 0 0 0 0 0 00010000
        0 0 0 0 0 0 0 0 0 0LE 0
PROTO  0 0 0 0 -45 0 0 0 0 0 00001000
        0 0 0 0 0 0 0 0 0 0LE 0
PROBA  0 0 0 0 0 -38 0 0 0 0 00000100
        0 0 0 0 0 0 0 0 0 0LE 0
PROGO  0 0 0 0 0 0 -16238 0 0 0 00000010
        0 0 0 0 0 0 0 0 0 0LE 0
PROMG  0 0 0 0 0 0 0 0 -9167 0 00000001
        0 0 0 0 0 0 0 0 0 0LE 0
VPTOT  0 0 0 0 0 0 0 0 0 0 -160 -0.67 -170 -180
        -380 -0.56 -1.6 1 0 0 0 0 0 0 0LE 0
AMORT  -46.8 -46.8 -46.8 -46.8 -46.8 -93.6 -93.6 -93.6 0 0 0 0 0 0 0
        0 1 0 0 0 0 0 0 0 0LE 0
CONAG  -5.13 -5.13 -4.93 -4.25 -3.86 -15.74 -13.52 -13.68 0 0 0 0 0 0 0
        0 0 1 0 0 0 0 0 0 0LE 0
VCONAG 0 0 0 0 0 0 0 0 0 0 00000000
        0 0 -20 1 0 0 0 0 0 0LE 0
DFUNRU 0 0 0 0 0 0 0 0 0 0 00000000
        -0.001 0 0 0 1 0 0 0 0 0LE 0
CCOMER 0 0 0 0 0 0 0 0 0 0 00000000
        -0.0067 0 0 0 0 1 0 0 0 0LE 0
CRED   485 485 457 733 1268 1476 1575 1416 5.50 0 0 0 0 0 0
        0 1 1 0 0 0 -1 -0.2 0LE 0
REMIN  3515 3515 749 3517 6832 12963 7518 13251 -5.50 0 0 0 0 0 0
        0 -1 0 -1 -1 -1 -0.08 0 0GE 9538
MARBRU 3515 3515 749 3517 6832 12963 7518 13251 -5.50 0 0 0 0 0 0
        0 -1 0 -1 -1 -1 -0.08 -1 0LE 0
DESP   485 485 457 733 1268 1476 1575 1416 5.50 0 0 0 0 0 0
        0 1 0 1 1 1 0.08 0 -1LE 1906
AMAXPE 0 0 0 0 0 1 1 1 1 0 00000000
        0 0 0 0 0 0 0 0 0 0LE 4
AMINMG 0 0 0 0 0 0 0 1 0 0 00000000
        0 0 0 0 0 0 0 0 0 0GE 2.1
AMAXME2 0 0 0 1 0 0 0 0 0 0 00000000
        0 0 0 0 0 0 0 0 0 0LE 1.5
AMINFE2 0 0 1 0 0 0 0 0 0 0 00000000
        0 0 0 0 0 0 0 0 0 0GE 0.5
AMINBA 0 0 0 0 0 1 0 0 0 0 00000000
        0 0 0 0 0 0 0 0 0 0GE 1

```

;

proc lp data=IRRIGA;

run;

```

/* Definição da macro para substituir os coeficientes aleatórios e
solucionar N vezes o problema de PL */
%macro AUX;
data RESULTS; stop; run; /*Cria um SAS dataset vazio(EMPTY) */
%do k=1 %to 1000 %by 1;
data PERIME P2;
set PERIME;
if _n_ eq 1 then output P2 ; else output PERIME;
run;
data LPAUX;
merge IRRIGA P2;
keep _id_
AB1 AB2 FE2 ME2 TO1 BA GO MG MOC PAB PFE PME PTO PBA PGO PMG
RB AMORT AGUA VAGUA FUR COM CRED MB DESP
_type_ _rhs_;
retain X1 _X2_ T1_ T2_ T3_ T4_ T5_ T6_ U1_ U2_ U3_ U4_ U5_ U6_ U7_ U8_;
if _n_ eq 1 then do; X1=X1; X2=X2; T1=T1; T2=T2; T3=T3; T4=T4; T5=T5;
T6=T6; U1=U1; U2=U2; U3=U3; U4=U4; U5=U5; U6=U6; U7=U7; U8=U8;
end;
X1=X1_; X2=X2_; T1=T1_; T2=T2_; T3=T3_; T4=T4_; T5=T5_; T6=T6_;
U1=U1_; U2=U2_; U3=U3_; U4=U4_; U5=U5_; U6=U6_; U7=U7_; U8=U8_;
if _id_ eq 'OBJET' then do;
AB1=X1; AB2=X2; FE2=T1; ME2=T2; TO1=T3; BA=T4; GO=T5; MG=T6;
end;
if _id_ eq 'CONAG' then do;
AB1=U1; AB2=U2; FE2=U3; ME2=U4; TO1=U5; BA=U6; GO=U7; MG=U8;
end;
run;
proc print;run;
proc LP data=LPAUX primalout=PP;

data TESTA;
set PP; keep Z X1 X2 X3 X4 X5 X6 X7 X8;
if _var_='OBJET' then Z=_value_;
if _var_='AB1' then X1=_value_;
if _var_='AB2' then X2=_value_;
if _var_='FE2' then X3=_value_;
if _var_='ME2' then X4=_value_;
if _var_='TO1' then X5=_value_;
if _var_='BA' then X6=_value_;
if _var_='GO' then X7=_value_;
if _var_='MG' then X8=_value_;

if Z='.' and X1='.' and X2='.' and X3='.' and X4='.'
and X5='.' and X6='.' and X7='.' and X8='.'
then delete;output;

```

```
proc means data=TESTA nway noprint missing;
var Z X1 X2 X3 X4 X5 X6 X7 X8;
output out=TESTAR
  max(Z X1 X2 X3 X4 X5 X6 X7 X8)=Z X1 X2 X3 X4 X5 X6 X7 X8;
run;
proc append base=EMPTY data=TESTAR;

run;
%end;
%mend aux;

%aux;

proc print data=TESTAR;run;

/* Impressão das soluções e da função objetivo */
proc print data=EMPTY;run;

/* Cálculo de algumas estatísticas simples da função objetivo e
das soluções do modelo de PL */
proc univariate plot normal data=EMPTY;
var Z X1 X2 X3 X4 X5 X6 X7 X8;
run;
```

APÊNDICE 2

AMOSTRA DOS DADOS OBTIDA DA SIMULAÇÃO PARA MARGEM BRUTA (Z) E PARA AS ATIVIDADES $X_1, X_2, X_3, X_4, X_5, X_6, X_7$ e

X_8 :

| | | | | | | | | |
|----------|---|-----|-----|-----|-----|-----|---|-----|
| 67732.13 | 0 | 2.0 | 0.5 | 0.0 | 2.5 | 1.0 | 0 | 3.0 |
| 69436.91 | 0 | 0.5 | 0.5 | 1.5 | 2.5 | 1.0 | 0 | 3.0 |
| 69047.64 | 0 | 2.0 | 0.5 | 0.0 | 2.5 | 1.9 | 0 | 2.1 |
| 74677.41 | 0 | 2.0 | 0.5 | 0.0 | 2.5 | 1.9 | 0 | 2.1 |
| 74458.35 | 0 | 2.0 | 0.5 | 0.0 | 2.5 | 1.9 | 0 | 2.1 |
| 69877.71 | 0 | 2.0 | 0.5 | 0.0 | 2.5 | 1.0 | 0 | 3.0 |
| 78157.29 | 0 | 0.5 | 0.5 | 1.5 | 2.5 | 1.0 | 0 | 2.1 |
| 75864.79 | 0 | 0.5 | 0.5 | 1.5 | 2.5 | 1.0 | 0 | 3.0 |
| 77431.71 | 0 | 0.5 | 0.5 | 1.5 | 2.5 | 1.0 | 0 | 3.0 |
| 78861.61 | 0 | 0.5 | 0.5 | 1.5 | 2.5 | 1.9 | 0 | 3.0 |
| 71243.71 | 0 | 0.5 | 0.5 | 1.5 | 2.5 | 1.0 | 0 | 3.0 |
| 78796.35 | 0 | 2.0 | 0.5 | 0.0 | 2.5 | 1.0 | 0 | 2.1 |
| 74580.68 | 0 | 2.0 | 0.5 | 0.0 | 2.5 | 1.0 | 0 | 3.0 |
| 82873.69 | 0 | 0.5 | 0.5 | 1.5 | 2.5 | 1.0 | 0 | 3.0 |
| 76740.50 | 0 | 2.0 | 0.5 | 0.0 | 2.5 | 1.9 | 0 | 3.0 |
| 67979.77 | 0 | 2.0 | 0.5 | 0.0 | 2.5 | 1.0 | 0 | 3.0 |
| 76304.22 | 0 | 0.5 | 0.5 | 1.5 | 2.5 | 1.0 | 0 | 3.0 |
| 63674.59 | 0 | 2.0 | 0.5 | 0.0 | 2.5 | 1.0 | 0 | 3.0 |
| 65292.28 | 0 | 2.0 | 0.5 | 0.0 | 2.5 | 1.0 | 0 | 3.0 |
| 77386.81 | 0 | 0.5 | 0.5 | 1.5 | 2.5 | 1.0 | 0 | 3.0 |
| 81968.01 | 0 | 2.0 | 0.5 | 0.0 | 2.5 | 1.0 | 0 | 3.0 |
| 63161.24 | 0 | 0.5 | 0.5 | 1.5 | 2.5 | 1.0 | 0 | 3.0 |
| 68706.28 | 0 | 2.0 | 0.5 | 0.0 | 2.5 | 1.0 | 0 | 3.0 |
| 73888.62 | 0 | 2.0 | 0.5 | 0.0 | 2.5 | 1.0 | 0 | 3.0 |
| 60396.56 | 0 | 0.5 | 0.5 | 1.5 | 2.5 | 1.0 | 0 | 3.0 |
