

MILTON PERCEUS SANTOS DE MELO

**PROGRAMAÇÃO LINEAR E SIMULAÇÃO MULTIDIMENSIONAL NO
MERCADO FINANCEIRO E *COMMODITIES***

RECIFE-PE – MARÇO/2012



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO

PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO

PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

**PROGRAMAÇÃO LINEAR E SIMULAÇÃO MULTIDIMENSIONAL NO
MERCADO FINANCEIRO E *COMMODITIES***

Dissertação apresentada ao Programa de Pós-Graduação em Biometria e Estatística Aplicada como exigência parcial à obtenção do título de Mestre.

**Área de Concentração: Modelagem
Estatística e Computacional**

Orientador: Prof. Dr. Moacyr Cunha Filho

Coorientador: Prof. Dr. Carlos Tadeu dos Santos Dias

RECIFE-PE – MARÇO/2012.

UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO

PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO

PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

**PROGRAMAÇÃO LINEAR E SIMULAÇÃO MULTIDIMENSIONAL NO
MERCADO FINANCEIRO E *COMMODITIES***

MILTON PERCEUS SANTOS DE MELO

Dissertação julgada adequada para obtenção do título de mestre em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 09/03/2012 pela Comissão Examinadora.

Orientador:

Prof. Dr. Moacyr Cunha Filho

Universidade Federal Rural de Pernambuco - UFRPE

Banca Examinadora:

Prof. Dr. Carlos Tadeu dos Santos Dias

Escola Superior de Agricultura "Luiz de Queiroz" / Universidade de São Paulo - ESALQ / USP

Prof. Dr. Tiago Alessandro Espínola Ferreira

Universidade Federal Rural de Pernambuco – UFRPE

Prof.^a. Dr.^a Tatijana Stosic

Universidade Federal Rural de Pernambuco - UFRPE

Agradecimentos

Agradeço primeiramente a Deus, por ter me guiado em todos os momentos e por ter me dado paciência e perseverança para superar todas as adversidades encontradas nesta grande jornada.

Pela orientação segura, confiança e incentivo, expresso meus maiores agradecimentos ao Prof. Dr. Moacyr Cunha Filho.

Pela atenção, confiança e eficiente orientação proporcionada no período de estágio na ESALQ/USP e em todo o desenvolvimento da dissertação, agradeço profundamente ao Prof. Dr. Carlos Tadeu Dos Santos Dias.

Agradeço ao Prof. Dr. Tiago Alessandro Espínola Ferreira pela preciosa colaboração e importantes sugestões apresentadas.

A Prof^a Dr.^a Tatijana Stosic pelas importantes sugestões apresentadas.

Aos colegas Dennis Marinho Oliveira R. de Souza, Djalma Beltrão da Costa Farias, pela amizade, colaboração e apoio.

Aos demais colegas, professores e funcionários do Departamento de Pós-Graduação em Biometria e Estatística aplicada da UFRPE, assim como a todos do Departamento de Estatística e Experimentação Agronômica da ESALQ/USP que, direta ou indiretamente contribuíram na realização do curso.

E por último e não menos importante agradeço a minha família, por ter me apoiado, incentivado e priorizado meus estudos em todos os momentos de minha vida.

Se você quiser ser bem sucedido, isso é simples: saiba o que você está fazendo,
ame o que você está fazendo e acredite no que você está fazendo.

Will Roger

Dedicatória

Dedico esta dissertação a minha família, em especial ao meu filho Thales Henrique Bezerra de Melo.

Resumo

Comumente após o plantio, a propriedade agrícola tenta avaliar o desempenho operacional dos anos anteriores através de estudos para posteriormente enfrentar um novo período de atividades com o início de uma nova safra. Os resultados levam os proprietários agrícolas a criar expectativas em torno de expansão ou redução de culturas, contratação de pessoal, compra de insumos e a novos investimentos. Todas estas indagações são pertinentes, provenientes de variações climáticas, manifestações de pragas e oscilações na economia que influenciam diretamente no preço dos produtos agrícolas. Contudo, os produtores deparam-se com a necessidade de tomar decisões que, muitas vezes, alteram completamente o processo operacional da produção. Em larga escala, estas decisões influenciam diretamente na oferta e no preço dos produtos. Entendendo que as técnicas utilizadas no planejamento agrícola são de grande valia para os produtores, uma metodologia de auxílio à decisão que consideram os riscos presentes no processo é aplicada no mercado à vista de *commodities* agrícolas com o objetivo de fornecer a melhor combinação de variáveis visando à otimização do lucro. Portanto, nesta dissertação foi utilizada uma técnica de simulação e programação linear, conjuntamente, levando em consideração as correlações existentes entre os coeficientes do modelo. Os dados foram obtidos de séries históricas referentes ao preço médio mensal de 44 *commodities* agrícolas comercializadas no mercado europeu à vista no período de janeiro de 1960 a setembro de 2011. Quanto à parte computacional, foi utilizado o software SAS, versão 9.2. Com a aplicação do método proposto, os resultados encontrados mostram que o uso da simulação sem dependência não subestimou o risco quando comparado à simulação com dependência e que a técnica não possibilitou a percepção do melhor momento para realizar o investimento, no entanto plenamente satisfatória como ferramenta de suporte à decisão.

Palavras-chave: Programação linear, simulação multidimensional, tomada de decisão, otimização, mercado financeiro.

Abstract

Usually after planting, the agricultural property attempts to assess the operational performance of the previous years through studies in order to face posteriorly a new period of activities with the beginning of a new crop. The results led the landowners to create expectations about the expansion or reduction of crops, staff hiring, supply purchasing and new investments. All these questions are relevant deriving from climatic variations, pest manifestations and from fluctuations in the economy which directly influence in the the prices of the agricultural products. However the producers face the need of making decisions which often completely change the operating process of production. In large scale these decisions influence directly the supply and the price of the products. Having in consideration that the techniques used in agricultural planning are of great value for the producers, a decision aid methodology which considers the risks in the process is applied on the spot market of agricultural *commodities* in order to provide the best combination of variables aiming the profit optimization. Therefore, in this it was jointly used simulation technique and linear programming taking into account the existent correlations between the coefficients of the model. The data were obtained from historic trends regarding the monthly average price of 44 agricultural *commodities* commercialized in the European spot market since January 1960 till September 2011. Regarding the computational part, it was used the software SAS, version 5.2. By applying the proposed method the results show that the use of simulation without dependence did not underestimate the risk when compared to the simulation with dependence and they also show that the technique did not allow the perception of the best moment to make the investment, yet fully satisfactory as a decision support tool.

Key words: Linear programming, multidimensional simulation, decision making, optimization, financial market.

Lista de Figuras

Figura 1 - Dominância Estocástica de 1º grau.....	23
Figura 2 - Dominância Estocástica 2º grau	24
Figura 3 - Fronteira Eficiente na análise de E-V	25
Figura 4 - Diagrama do modelo de Programação Linear.....	33
Figura 5 - Gráfico da densidade triangular Triang (a, b, m)	40
Figura 6 - Ilustração do método da transformação inversa para variáveis aleatórias contínuas	44
Figura 7 - Histograma da variável Café IOC gerada com distribuição triangular	49
Figura 8 - Histograma da variável Couro Bovino gerada com distribuição triangular	49
Figura 9 - Histograma da variável Óleo Algodão gerada com distribuição triangular	50
Figura 10 - Histograma da variável Grãos de Cacau gerada com distribuição triangular	50
Figura 11 - Histograma da variável Madeira Serrada gerada com distribuição triangular	51
Figura 12 - Histograma da variável Chá gerada com distribuição triangular	51
Figura 13 - Histograma da variável Juta gerada com distribuição triangular.....	52
Figura 14 - Histograma da margem bruta com dependência	54
Figura 15 - Histograma da margem bruta sem dependência.....	54

Lista de Tabelas

Tabela 1 - Parâmetros de mínimo, máximo e moda utilizados na geração de números aleatórios com distribuição triangular	45
Tabela 2 - Coeficiente de correlação de Pearson para as variáveis geradas.....	48
Tabela 3 – Taxa média de alocação de recursos	53

Sumário

1	INTRODUÇÃO	14
1.1	OBJETIVOS.....	15
1.2	ESTRUTURA DA DISSERTAÇÃO	15
2	REVISÃO DE LITERATURA	17
2.1	ANÁLISE HISTÓRICA DA TEORIA DA DECISÃO NA AGRICULTURA.....	17
2.2	TEORIA DA DECISÃO E ATIVIDADE AGRÍCOLA	19
2.2.1	DEFINIÇÃO DE RISCO E INCERTEZA NA AGRICULTURA.....	20
2.3	MODELOS DE DECISÃO COM INCORPORAÇÃO DE RISCO	21
2.3.1	DOMINÂNCIA ESTOCÁSTICA	21
2.3.2	PROGRAMAÇÃO QUADRÁTICA DE RISCO	25
2.3.3	MOTAD.....	27
2.3.4	TARGET - MOTAD	29
2.3.5	PROGRAMAÇÃO COM RESTRIÇÃO DO TIPO FOCO-PERDA.....	30
2.3.6	PROGRAMAÇÃO LINEAR E SIMULAÇÃO MULTIDIMENSIONAL.....	32
2.3.6.1	PROGRAMAÇÃO LINEAR	32
2.3.6.2	OPERACIONALIZAÇÃO DO MODELO	34
3	MATERIAIS E MÉTODOS	35

3.1	A TEORIA DE SELEÇÃO DE PORTFÓLIO E O MERCADO FINANCEIRO	35
3.2	MASSA DE DADOS	36
3.3	METODOLOGIA PROPOSTA.....	36
3.3.1	SIMULAÇÃO DE MONTE CARLO	37
3.3.1.1	IDENTIFICAÇÃO DAS DISTRIBUIÇÕES DAS VARIÁVEIS	38
3.3.1.2	GERAÇÃO DE NÚMEROS ALEATÓRIOS.....	40
3.3.1.2.1	GERAÇÃO DE NÚMEROS ALEATÓRIOS COM DISTRIBUIÇÃO NORMAL	41
3.3.1.2.2	GERAÇÃO DE NÚMEROS ALEATÓRIOS COM DISTRIBUIÇÃO UNIFORME	43
3.3.1.2.3	GERAÇÃO DE NÚMEROS ALEATÓRIOS COM DISTRIBUIÇÃO TRIANGULAR	45
3.3.1.2.4	FORMULAÇÃO MATEMÁTICA PARA A GERAÇÃO DE NÚMEROS ALEATÓRIOS A PARTIR DO MODELO DE MARKOWITZ	46
4	RESULTADOS E DISCUSSÃO.....	48
4.1	RESULTADOS	48
4.2	CONCLUSÕES E SUGESTÕES.....	54
	REFERÊNCIAS	56
	APÊNDICE A: COEFICIENTE DE CORRELAÇÃO ENTRE AS <i>COMMODITIES</i> ESCOLHIDAS.....	61
	APÊNDICE B: PROGRAMA PARA SIMULAR COEFICIENTES DEPENDENTES E SOLUCIONAR O PROBLEMA DE PROGRAMAÇÃO LINEAR	61
	APÊNDICE C: PROGRAMA PARA SIMULAR COEFICIENTES INDEPENDENTES E SOLUCIONAR O PROBLEMA DE PROGRAMAÇÃO LINEAR	65

1 INTRODUÇÃO

Em qualquer área da ciência, o interesse por novas técnicas manifesta-se frequentemente quando se procuram soluções para problemas de difícil tratamento ou análise de relações entre variáveis.

Em vários casos, a matemática nos oferece a estrutura ideal para a construção e representação das partes que compõe o problema. De fato, nos oferece uma poderosa linguagem para a formalização, representação e análise das relações e conhecimentos adquiridos como fruto do trabalho de investigação. Especialmente em situações que envolvem mudanças ao longo do tempo e processos dinâmicos em geral. Como a estatística é um ramo de matemática aplicada, os seus métodos são rigorosos e precisos. Apesar da objetividade que a matemática confere aos métodos estatísticos, deve ter-se em conta que os seus resultados incorporam subjetividade. Tal subjetividade resulta principalmente da qualidade das medidas e da variabilidade de diversos fenômenos que influenciam nas observações (DASKALAKI, 2011).

Todo investidor que almeja sucesso no mercado de ativos deve combinar educação financeira, pleno conhecimento do mercado e, acima de tudo, uma visão de longo prazo. Essas são ferramentas fundamentais e de grande valia para quem deseja ter sucesso no mercado de ações (DRIESSEN et al., 2007).

Os quatro capítulos deste trabalho compreendem a análise de ferramentas de auxílio à tomada de decisão no planejamento agrícola e sua utilização no mercado financeiro.

A abordagem quantitativa, e conseqüentemente matemática e estatística, nos permite formalizar as relações que regem a dinâmica de seleção de ativos financeiros e a análise de comportamento a partir de simulações para nos dar subsídio para a tomada de decisão e percepção de risco.

Este trabalho dá ênfase especial a modelos de auxílio à tomada de decisão no planejamento agrícola e sua aplicação no mercado financeiro, em especial o mercado à vista de *commodities* agrícolas.

1.1 OBJETIVOS

Esta dissertação tem como objetivo principal utilizar uma tecnologia agrícola de auxílio à decisão no mercado financeiro de *commodities* agrícolas.

Com a coleta de dados históricos, pretende-se verificar quais distribuições de probabilidade fornecem um bom ajuste para aplicar técnicas de simulação e geração de números aleatórios com e sem dependência entre os coeficientes estocásticos.

Podemos destacar os seguintes objetivos específicos:

- Gerar números aleatórios bem comportados de acordo com a distribuição de origem;
- Criar números aleatórios a partir do modelo de Markowitz (MARKOWITZ, 1952);
- Comparar os resultados a partir da geração de números aleatórios independentes e correlacionados;
- Realizar uma análise dos resultados para a tomada de decisão e risco.

1.2 ESTRUTURA DA DISSERTAÇÃO

O trabalho está dividido em quatro capítulos.

O segundo capítulo apresenta uma revisão de literatura com uma análise histórica de modelos e métodos utilizados no processo de decisão, assim como os aspectos conceituais dos principais modelos de auxílio à decisão que consideram o risco no planejamento agrícola.

No terceiro capítulo mostramos a ligação existente entre os modelos de decisão no planejamento agrícola e o mercado financeiro, além de tratar da especificação dos dados, procedimentos e estruturação do modelo utilizado.

No quarto capítulo apresentamos os resultados das análises efetuadas a partir do processo de simulação.

Nos apêndices são apresentados as linhas de programação referentes a simulação de coeficientes dependentes e correlacionados para solucionar o problema de programação linear.

2 REVISÃO DE LITERATURA

Nas primeiras seções deste capítulo apresentam-se uma análise histórica dos principais modelos de decisão que nortearam o planejamento agrícola assim como um enfoque da teoria de decisão na atividade agrícola e a diferenciação sobre os conceitos de risco e incerteza na agricultura. Uma abordagem sobre os principais modelos de tomada de decisão que consideram os riscos pertinentes no processo de planejamento agrícola são vistos na seção seguinte.

2.1 ANÁLISE HISTÓRICA DA TEORIA DA DECISÃO NA AGRICULTURA

No século XV, a partir do renascimento, houve um despertar pelo conhecimento. Com os avanços estabelecidos por Nicolau Copérnico, Johannes Kepler e Galileu Galilei houve uma crescente por parte dos matemáticos da época, considerando que a matemática teria muito a contribuir para o entendimento do mundo real (CUSINATO et al., 2003).

No século XVII, Blaise Pascal e Pierre de Fermat desenvolveram os primeiros passos da teoria das probabilidades quando tentavam explicar o comportamento humano. Em 1731 Nicolas Bernoulli propôs um exemplo, conhecido como Paradoxo de São Petersburgo que podia contestar a hipótese de Pascal-Fermat. A solução deste paradoxo foi proposta pelo seu sobrinho, Daniel Bernoulli 1738.

Para Cusinato et al (2003) além da solução de Daniel Bernoulli ter sido um marco inicial para a teoria da utilidade esperada, ela foi esquecida por um longo tempo. Gabriel Cramer forneceu uma solução semelhante através de carta enviada a Nicolas Bernoulli. Em um pós-escrito ao artigo de 1738, Daniel Bernoulli reconhece o trabalho de Cramer citando-o em várias partes.

Ainda no século XVIII Jeremy Bentham redescobriu o conceito da função utilidade através de uma filosofia utilitarista, escola filosófica que estabelece a prática de suas ações de acordo com a utilidade, baseando-se em preceitos éticos. Para Cusinato et al. (2005) esta foi a descoberta mais influente da utilidade esperada.

Na década de 1870, marginalistas introduzem o conceito de utilidade na teoria econômica modelando contextos sem incorporação de incerteza.

Contudo, Dias (1996) esclarece que nos séculos XVIII, XIX e início do século XX a teoria da utilidade se desenvolveu de forma confusa, pois conceitos tais como funções diferença de valor, funções de valores ordinais e funções-utilidade foram confundidas.

Em 1906 Vilfredo Pareto introduz a abordagem ordinal da teoria da utilidade, na década de 1930 John Hicks e R.G.D. Allen apresentam a abordagem “operacionalista” da teoria da utilidade.

Na década de 40, o cientista norte-americano George B. Dantzig, oferecendo subsídios técnicos para a tomada de decisão na segunda guerra mundial, desenvolveu o método simplex, método matricial para resolver problema de programação linear, e publicou posteriormente no artigo “Linear programming and extensions” de 1963 (CAIXETA-FILHO, 2004).

Em 1940, durante a 2ª Guerra, Von Neumann e Ulam utilizaram o método de Monte Carlo quando trabalhavam na solução de problemas relacionados com a construção de armas nucleares. Mais foi 1949 que Metropolis e Ulam, deram definitivamente origem ao método de Monte Carlo (METROPOLIS et al., 1949).

Markowitz desenvolveu a teoria do portfólio que relaciona o retorno do investimento e o risco que se corre a optar por tal investimento. Considera a construção de uma carteira de investimentos com lucro propício pode ser formada por ativos pouco ou negativamente correlacionados (MARKOWITZ, 1952). Para Choy (2011), tal artigo foi um marco na formulação e no desenvolvimento da teoria de diversificação de investimentos sob condições de risco.

Markowitz (1959) coloca os fundamentos para a introdução e aplicação da programação estocástica em problemas econômicos representando assim um avanço no processo de decisão frente ao risco. Neste mesmo artigo foi dado um passo importante no sentido de restringir as generalidades do teorema da utilidade esperada com a análise de média-variância, proposto em 1952 por Markowitz, que utiliza os dois primeiros momentos da distribuição de probabilidades dos retornos ou da renda (CHOY, 2011; MARKOWITZ, 1959).

Pratt (1964) introduziu algumas medidas de aversão ao risco na econometria que mais tarde ficaram conhecidas como coeficiente de aversão de Arrow-Pratt. Os coeficientes também levaram o nome de Kennett Arrow porque Arrow já havia esboçado algumas análises sobre o tema quando Pratt as tomou como medida de aversão ao risco. Em seu artigo ele reconhece a contribuição de Arrow. Para Cusinato (2003), este artigo continua sendo referência importante no campo de medidas de aversão ao risco.

Já em 1967, Boussard e Petit desenvolveram um Modelo de Programação Linear do Tipo Foco-Perda que se baseia na hipótese de que os agricultores escolhem, entre alternativas viáveis, aquele conjunto de soluções que maximiza sua renda esperada, na condição de que a possibilidade de falência tenha uma probabilidade muito baixa de ocorrer (BOUSSARD E PETIT, 1967).

Hazell desenvolveu a aplicação de Programação Quadrática à agricultura em 1968. Posteriormente Hazell (1971), reconhecendo as dificuldades da aplicação do método, propôs o uso do Motad - Minimization of Total Absolute Deviation (HAZELL, 1971).

Em 1974, foi apresentado os critérios de Dominância Estocástica, instrumento metodológico que envolve toda a distribuição cumulativa de probabilidade dos retornos e permite identificar alternativas mais eficientes através de comparações entre as mesmas (ANDERSON, 1974).

Já em 1983, Tauer desenvolveu um modelo de programação matemática alternativo ao modelo MOTAD, o Target-MOTAD, que apresenta vantagens de se gerar soluções eficientes no sentido de dominância estocástica de 2º grau, ou seja, ideal para agricultores avessos ao risco (TAUER, 1983).

Com o propósito de se aproximar da realidade, Dias propôs um método de inclusão de risco em modelos de planejamento agrícola através de técnicas de simulação e programação linear, conjuntamente, levando em consideração relações de dependência existente entre os coeficientes (DIAS, 1996).

2.2 TEORIA DA DECISÃO E ATIVIDADE AGRÍCOLA

Diante de um contexto baseado em transformações a agricultura modificou-se e o produtor rural passou a necessitar de conhecimento e de agilidade na busca de competitividade e, até mesmo, de sobrevivência neste cenário de complexidade.

Similarmente, a manifestação de pragas e as seguidas adversidades climáticas forçam a evolução constante de novos modelos de gestão e organização rural.

Segundo Peche Filho (2007) o produtor no gerenciamento agrícola depara-se constantemente com a necessidade de tomar decisões, e em alguns casos, essas decisões alteram completamente o processo operacional da produção. O exemplo clássico é na escolha de técnicas para atender as muitas

variáveis que envolvem o preparo do solo, um estudo é necessário para definir as alternativas melhor indicadas aos fatores da propriedade agrícola.

Visando auxiliar tal processo de decisão recorre-se, aos princípios da teoria de decisão. Pressupõe-se que seja possível especificar convenientemente todas as alternativas de ação possíveis, os estados da natureza e principalmente fatores que em conjunto irão influenciar os resultados de se tomar e implementar uma decisão, considerando suas probabilidades de ocorrência, as consequências de se implementar cada ação para os estados da natureza considerados e uma função de utilidade a ser maximizada (DIAS, 1996).

Os riscos e incertezas presentes na produção agrícola têm um papel importante no processo de decisão do agricultor a cerca da alocação de seus recursos, por esta razão, o agricultor é forçado a tomar decisões sobre as quais não tem o conhecimento completo das informações ((PARIKH; SHAH, 1994).).

Existe uma literatura abundante que utiliza diversos modelos de tomada de decisão, como por exemplo, Dominância Estocástica, Programação Quadrática, Motad, Programação tipo Foco-Perda, que levam em consideração os riscos e incertezas presentes no processo (BOUSSAR; PETIT, 1967; CRUZ, 1984; MATTUELLA, L.J. 1990; MOSS, 2010; MOUTINHO et al.,1978; PÉRES, 1984; SOARES, F.B. 1977; TURVEY, C.G. et al., 2005).

2.2.1 DEFINIÇÃO DE RISCO E INCERTEZA NA AGRICULTURA

A agricultura constitui uma atividade econômica bastante complexa, dado o caráter aleatório de vários fenômenos climáticos, biológicos e o grande número de variáveis que afetam as oportunidades de comercialização e lucro.

No ambiente de decisão da agricultura as variáveis risco e incerteza têm sido largamente utilizadas, onde o risco é caracterizado por situações que ocorrem com distribuição de probabilidade conhecida e a incerteza por situações onde a probabilidade de ocorrência não pode ser antecipada (BITTENCOURT et al.,1996).

Para Cruz (1984), essa diferenciação clássica entre situações de risco e incerteza, apresentadas a princípio por Knight (1921) e citada por Bittencourt & Sampaio (1998), não é apropriada para a agricultura pois o agricultor possui uma idéia subjetiva da probabilidade de ocorrência de um evento e não uma probabilidade objetiva, como pensava Knight (1921), uma vez que na agricultura a disponibilidade de dados não permite um tratamento assintótico para os dados. Quanto às situações de incerteza, estas seriam caracterizadas por total desinformação do agricultor, coisa que não é aplicável.

Desta forma, o conceito de incerteza coincide com o de risco com probabilidade subjetiva, podendo os termos ser usados indistintamente.

2.3 MODELOS DE DECISÃO COM INCORPORAÇÃO DE RISCO

Existem na literatura diversos trabalhos que utilizam modelos de programação que incluem os riscos e incertezas presentes no auxílio à decisão.

Dillon em 1975, assim como Bittencourt e Sampaio em 1998, faz uma análise teórica dos principais modelos de decisão que incorporam incerteza mostrando vantagens e desvantagens de cada modelo.

Já em 1996, Dias faz uma análise de risco a partir de simulações realizadas pelo método de Monte Carlo, estudando a distribuição de frequências dos valores ótimos obtidos após as simulações e realizando uma análise comparativa dos resultados que consideram as correlações existentes entre as variáveis com os resultados que não consideram estas correlações.

No âmbito internacional, autores utilizam e valorizam este modelos, Turvey et all (2005) realiza uma abordagem de técnicas de análise de portfólio centralizando o trabalho nos modelos Motad, Média e Variância de Markowitz de 1952 e 1959 e no Target Motad de Tauer de 1983, já Salimonu et all (2008) faz uma aplicação do Target-Motad sugerindo planos de alocação de recursos para agricultores da Nigéria, Sharma et all (2010) faz uma análise de risco usando modelos de programação não-linear baseado em diferentes cenários e Moss (2010) ressalva que o livro por ele escrito é primeira importante pesquisa de análise de risco realizada desde a análise de decisão publicada por Anderson (1974).

Destacamos os modelos Dominância Estocástica, Programação Quadrática de Risco, Motad, Target Motad, Programação com Restrição tipo Foco-Perda e Programação linear e simulação multidimensional por estarem mais presentes na literatura consultada.

2.3.1 DOMINÂNCIA ESTOCÁSTICA

O método de Dominância Estocástica foi apresentado em 1974 por Anderson para variáveis contínuas, sendo que o mesmo pode ser aplicado para variáveis discretas. Teve como proposta desenvolver técnicas de

dominância estocástica, a partir das preferências do tomador de decisão que são expressas pela Função Utilidade, que levam em conta toda a distribuição cumulativa de probabilidade dos retornos ao invés de simplesmente a média e a variância (ANDERSON, 1974).

Uma série de regras permite a seleção das combinações por este método, e esta é feita em três estágios: Primeiro, Segundo e Terceiro Grau de Dominância Estocástica (ANDERSON, 1974; ROMAN, 2006).

Se a Função Utilidade é função de apenas uma variável e temos duas distribuições $f(x)$ e $g(x)$, que são definidas no intervalo $a \leq x \leq b$ onde $f(x)$ é referente à alternativa F e $g(x)$ é referente à alternativa G .

Uma vez definidas as funções de probabilidade elas são convertidas em funções de distribuição acumulada para valores de R entre a e b . Conhecendo-se as funções de probabilidade podem-se aplicar as regras de dominância estocástica de primeiro grau utilizando-se as fórmulas abaixo:

$$F_1(r) = \int_a^r f(x) dx \qquad G_1(r) = \int_a^r g(x) dx$$

A Dominância Estocástica de Primeiro grau de F sobre G ocorre, se:

$$F_1(r) \leq G_1(r)$$

Denotando a Função Utilidade por $U(x)$, de forma que se apresente crescente, $\frac{\partial U(x)}{\partial x} > 0$, o agricultor irá preferir as distribuições que sejam dominantes em primeiro grau.

A Figura 1 representa uma análise gráfica de distribuições cumulativas de quatro opções de investimento, a , b , c e d . É nítida a dominância de a sobre b , já que $F_a(R)$ está totalmente à direita de $F_b(R)$ de forma que $F_a(R) \leq F_b(R)$. Da mesma forma c e d são preferíveis a b , pois $F_c(R) \leq F_b(R)$ e $F_d(R) \leq F_b(R)$ para qualquer valor de R .

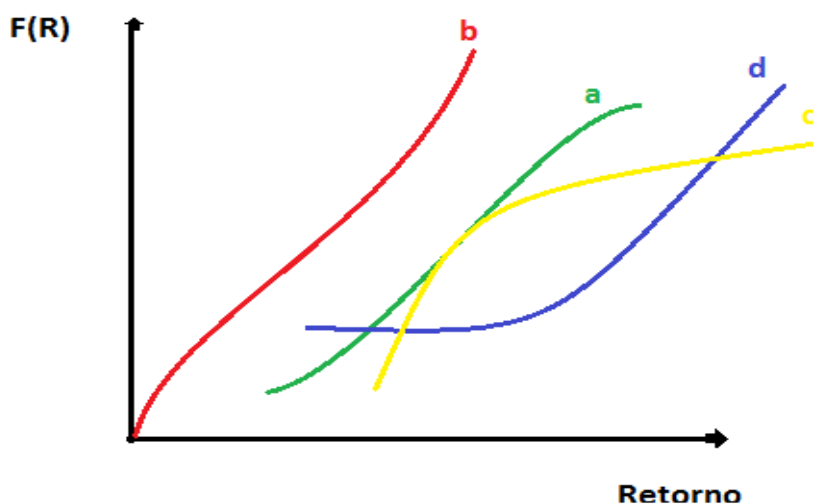


Figura 1 - Dominância Estocástica de 1º grau

Apesar de possuírem uma interseção, *c* domina *a* já que não se cruzam e *c* está à direita de *a*. Já a distribuição *d* cruza tanto com *a* como com *c*, desta forma não podemos realizar nenhuma afirmação.

Pelo critério de Dominância Estocástica de Primeiro grau a condição necessária e suficiente para que uma domine a outra é que duas funções cumulativas de probabilidade não podem se interceptar (ANDERSON, 1974).

Muitas vezes correm situações, como a apresentada na Figura 2 entre *d-a* e *d-c* em que em que não podemos identificar a alternativa mais eficiente pelo Primeiro Grau de Dominância Estocástica, emprega-se então de Segundo Grau de Dominância Estocástica (MOUTINHO, 1978).

Se sua função utilidade, além de apresentar $\frac{\partial U(x)}{\partial x} > 0$, apresente também $\frac{\partial^2 U(x)}{\partial x^2} > 0$, mostra que o agricultor terá uma função utilidade crescente e, sendo avesso ao risco, irá preferir distribuições que sejam dominantes em segundo grau.

Então, dados

$$F_2(r) = \int_a^r F_1(x) dx \qquad G_2(r) = \int_a^r G_1(x) dx$$

a Dominância Estocástica de Segundo grau de *F* sobre *G* ocorre, se:

$$F_2(r) \leq G_2(r)$$

A Figura 2 representa graficamente G e F (um sinal + representa os intervalos em que $F(R) < G(R)$ e um sinal – o intervalo em que $F(R) > G(R)$).

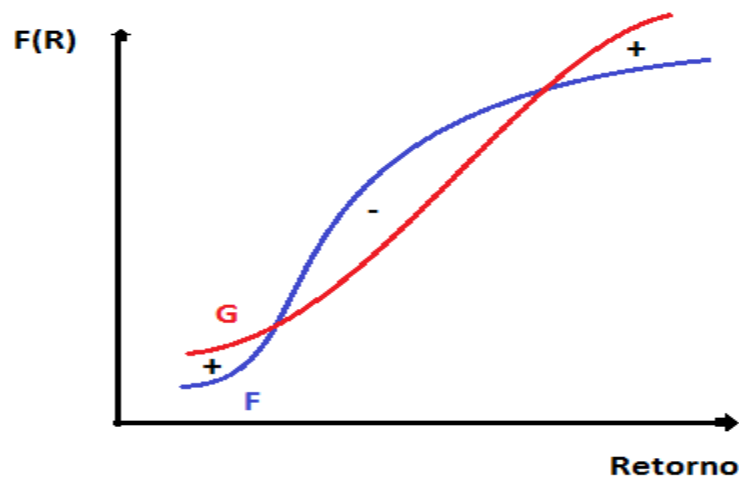


Figura 2 - Dominância Estocástica 2º grau

Uma vez que para todo retorno possível à área cumulativa entre as duas distribuições é positiva, o critério de dominância estocástica de 2º grau é satisfeito e F domina G para os indivíduos avessos ao risco.

Frequentemente ocorrem situações onde a Função Utilidade, além de apresentar $\frac{\partial^2 U(x)}{\partial x^2} > 0$, apresente também $\frac{\partial^3 U(x)}{\partial x^3} > 0$, logo o tomador de decisão deverá preferir as distribuições que sejam dominantes em terceiro grau por além de ter a função de utilidade crescente e ser avesso ao risco, ele terá também uma aversão decrescente ao risco (MOUTINHO, 1978; ROMAN, 2006). Este critério de dominância assume que além de ter utilidade crescente e ser avesso ao risco o investidor possui aversão decrescente ao risco.

Onde a Dominância Estocástica de Terceiro Grau segue as mesmas regras anteriores, sendo:

$$F_3(r) = \int_a^r F_2(x) dx \qquad G_3(r) = \int_a^r G_2(x) dx$$

Ocorre Dominância Estocástica de Terceiro Grau de H sobre G se:

$$H_2(r) \leq G_2(r)$$

A prova está contida em Anderson (1974), detalhes sobre os critérios de dominância estocástica, aplicados a variáveis discretas podem ser vistos em Cruz (1984), Tauer (1983) e Cusinato (2005).

2.3.2 PROGRAMAÇÃO QUADRÁTICA DE RISCO

A Programação Quadrática de Risco é usada para se obter um planejamento agrícola em condições de incerteza permitindo considerar a estocasticidade da produção agrícola, isto é, a variação dos rendimentos ou retornos das atividades onde o risco é medido pela variância dos retornos esperados (CRUZ, 1984; DIAS, 1996; DILLON, 1975; MARKOWITZ, 1952; MARKOWITZ, 1959; MOSS, 2010; PÉRES, 1981; TURVEY, 2005).

Presume que o tomador de decisão determina que os portfólios eficientes sejam aqueles que apresentam menor variância para um mesmo retorno esperado, ou a alternativa que apresenta maior retorno esperado para um mesmo nível de variância.

O objetivo da Programação Quadrática é desenvolver um conjunto de planos viáveis com a propriedade de mínima variância (risco) para um determinado nível de renda esperado. Tais planos são conhecidos como pontos eficientes E-V (média e variância) e definem a fronteira eficiente para todos os planos viáveis. A aceitação de qualquer uma das soluções ou planos dependerá das preferências do tomador de decisão, de acordo com a sua função de utilidade que é a função que representa o grau de satisfação do agricultor (CUSINATO, 2003; MARKOWITZ, 1959; MOSS, 2010).

Com o conjunto de planos eficientes definido, o decisor escolherá aquele que melhor representa a sua preferência entre as várias rendas esperadas e suas variâncias relacionadas, definidas pela função E-V. Como podemos observar na Figura 3, quando essa função utilidade pode ser definida, um único plano pode ser considerado como de máxima utilidade.

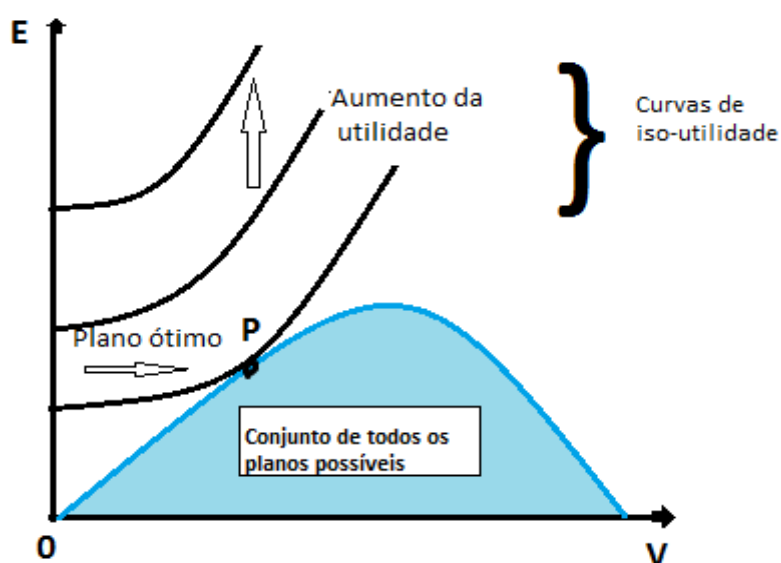


Figura 3 - Fronteira Eficiente na análise de E-V

As curvas de iso-utilidade são utilizadas para representar as preferências do investidor frente ao risco e ao retorno. Para o caso de um investidor racional (avesso ao risco), suas preferências serão maiores para um maior retorno esperado, conforme cresce o risco do investimento. Assumindo que incrementos adicionais no risco demandam incrementos ainda maiores no retorno esperado, as curvas de iso-utilidade serão convexas.

O modelo de Programação Quadrática pode ser então definido como:

Minimizar:

$$V = \sum_{j=1}^n \sum_{k=1}^n \sigma_{jk} x_j x_k$$

Sujeito a:

$$\checkmark \sum_{j=1}^n c_j x_j = Z \quad (\beta = 0, 1, \dots, Z^*) \quad (1)$$

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

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

Em que:

- V Representa a variância;
- σ_{jk} Representa a covariância da margem bruta entre a j -ésima e a k -ésima atividade quando $j \neq k$ e variância da margem bruta da j -ésima atividade quando $j = k$;
- x_j Representa a j -ésima atividade agrícola;
- a_{ij} Representa a exigência técnica da i -ésima restrição por unidade de j -ésima atividade agrícola;
- c_j Margem bruta média por unidade da j -ésima atividade agrícola;
- b_i Um escalar que representa o i -ésimo recurso disponível;
- Z Parâmetro representando os níveis escolhidos de renda;
- Z^* Margem bruta total resultante da estratégia ótima indicada pelo modelo de Programação Linear sem risco.

Parametrizando Z de 0 a Z^* , obtém-se uma sequência de estratégias com margens brutas totais crescentes e variâncias mínimas.

Segundo Cruz (1984) quando uma alternativa A comparada com uma alternativa B apresenta a maior média e maior variância, pode se dizer que ambas são eficientes sob o critério de análise de E-V. Esta característica tende a ser indesejável, pois em certos casos, um alternativa A pode apresentar um retorno médio muito superior, e a penas um pequeno acréscimo de variância em relação a alternativa B será suficiente para tornar ambas as alternativas igualmente desejáveis.

A aplicação deste método é limitada pela necessidade de informações sobre as covariâncias entre os retornos líquidos, por ser aplicável apenas em situações onde a margem bruta se apresente normalmente distribuída e por se restringir aos casos onde apenas os coeficientes c_j da função objetivo são aleatórios (DILLON, 1975; MOSS, 2010).

2.3.3 MOTAD

O modelo Motad (Minimização dos Desvios Absolutos Totais) foi proposto por Hazell em 1971 a partir de modificações introduzidas no método apresentado por Markowitz em 1959.

A modificação sugerida apresenta a vantagem de substituir a variância pelos desvios absolutos com relação à média no uso da programação linear como parâmetro de decisão do indivíduo (HAZELL, 1971; MOSS, 2010).

Com este modelo, tem-se a vantagem da possibilidade de uso dos pacotes convencionais de programação linear e de se poder estabelecer uma fronteira eficiente comparável a da programação quadrática (TURVEY, 2005).

O desvio absoluto com relação à média A pode ser definido como segue:

$$A = \frac{1}{s} \sum_{h=1}^s \sum_{j=1}^n |c_{hj} - g_j| x_j$$

Em que:

- A Representa um estimador não tendencioso do desvio de renda média absoluta da população;
- s Representa o número de observações numa amostragem aleatória de margens brutas;
- c_{hj} Representa a margem da h-ésima observação da j-ésima atividade;
- g_j Representa a média amostral das margens brutas de atividade j ;
- x_j Representa o nível da j-ésima atividade.

Sendo A a medida de incerteza e considerando E como a média da margem bruta, têm-se os parâmetros fundamentais para a seleção das atividades.

Os planos eficientes são encontrados minimizando-se $E - A$ (DILLON, 1975; PÉRES, 1981).

Quando a série histórica de margens brutas consideradas se referir a um evento com distribuição normal, uma medida não tendenciosa da variância é obtida através de :

$$V = A^2 \frac{\pi S}{2(s-1)}$$

E desta forma a utilização de A como medida de risco seria justificada (HAZELL, 1971).

Sendo $\sum_{h=1}^s y_h^+$ como a soma de valores absolutos dos desvios positivos da margem bruta ao redor da média esperada, e $\sum_{h=1}^s y_h^-$ a soma de valores absolutos dos desvios negativos da margem bruta ao redor da média esperada e assumindo que o desvio positivo da média $\sum_{h=1}^s y_h^+$ seja igual ao desvio negativo $\sum_{h=1}^s y_h^-$, o modelo de programação linear pode ser resolvido pela minimização da soma dos valores absolutos dos desvios negativos das margens brutas totais em torno da média.

O modelo MOTAD proposto por Hazell (1971) é desenvolvido da seguinte forma:

Minimizar

$$\sum_{h=1}^s y_h^-$$

Sujeito a:

$$\checkmark \sum_{j=1}^n |c_{hj} - g_j| x_j + y_h^- \geq 0 \quad (h = 1, 2, \dots, s) \quad (1)$$

$$\checkmark \sum_{j=1}^n f_j x_j = \beta \quad (\beta = 1, 2, \dots, z^*) \quad (2)$$

$$\checkmark \sum_{j=1}^n a_{ij} x_j \leq b_i \quad (i = 1, 2, \dots, m) \quad (3)$$

$$\checkmark y_h^-, x_j \geq 0 \quad (h = 1, 2, \dots, s; i = 1, 2, \dots, m) \quad (4)$$

Em que:

- y_h^- Representa o valor absoluto do desvio negativo da margem bruta ao redor da média esperada;
- β Representa o coeficiente de parametrização;
- z^* Representa o valor máximo dado pela programação linear;
- f_j Representa a margem bruta esperada da j -ésima atividade;

- a_{ij} Representa o coeficiente técnico do i-ésimo recurso da j-ésima atividade;
- b_i Representa o nível de fatos limitante associado a i-ésima restrição;
- n Representa o número de atividades;
- m Representa o número de restrições.

Parametrizando-se o valor da margem bruta β desde 0 até z^* podemos gerar a fronteira eficiente (MOSS, 2010; TURVEY, 2005).

Segundo Cruz (1984), a escolha da solução eficiente é, portanto, definida pelo produtor de acordo com sua função utilidade.

Autores destacam como principais desvantagens deste método o fato de que sua validade pode estar restrita a casos onde é possível considerar aproximadamente normal à distribuição da margem bruta total e ao fato de se tratarem como fonte de incerteza apenas os dois primeiros momentos da distribuição da margem bruta (DILLON, 1975; MOSS, 2010; PÉRES, 1981, TURVEY, 2005).

2.3.4 Target - MOTAD

O Target-MOTAD é um modelo de programação matemática alternativo ao modelo MOTAD, foi apresentado com o propósito de se gerar soluções eficientes no sentido de dominância estocástica de 2º grau, ou seja, ideal para agricultores avessos ao risco (MARQUES, 2009; TAUER, 1983; TURVEY, 2005).

Sua operacionalização é dada da seguinte forma:

Maximizar:

$$E(z) = \sum_{j=1}^n c_j x_j$$

Sujeito a:

$$\checkmark \sum_{j=1}^n a_{kj} x_j \leq b_j \quad (k = 1, 2, 3, \dots, m) \quad (1)$$

$$\checkmark T - \sum_{j=1}^n c_{rj} x_j - y_r \leq 0 \quad (r = 1, 2, 3, \dots, s) \quad (2)$$

$$\checkmark \sum_{r=1}^s p_r y_r = \beta \quad \text{com } \beta = M \rightarrow 0 \quad (3)$$

$$\checkmark x_j \geq 0 \quad (4)$$

$$\checkmark y_r \geq 0 \quad (5)$$

Em que:

- $E(z)$ Representa o retorno esperado do plano de soluções ótimas;
- c_j Representa o retorno esperado da atividade j ;
- a_{kj} Representa o coeficiente técnico do k -ésimo recurso da j -ésima atividade;
- b_j Representa o recurso disponível para a j -ésima atividade;
- T Representa o nível de retorno desejado (alvo);
- c_{rj} Representa o retorno esperado da atividade j para o estado de natureza da observação r ;
- γ_r Representa o desvio entre o alvo T e o estado da natureza da observação r ;
- p_r Representa a probabilidade do estado da natureza de ocorrência observação r ;
- β Representa uma constante parametrizada que vai de M até 0 ;
- m Representa o número máximo equações de restrições e de recursos;
- s Representa o número de observações de estados da natureza.

Onde o risco é medido pela soma esperada dos desvios negativos resultantes com o retorno desejado (alvo) e o retorno é medido pelo produto da soma dos retornos esperados de cada atividade pelo seu respectivo nível individual, a prova está contida em Tauer (1983).

2.3.5 PROGRAMAÇÃO COM RESTRIÇÃO DO TIPO FOCO-PERDA

Segundo Moss (2010) modelo de programação linear do tipo Foco-Perda foi proposto por Shackle (1949 - 1961) e aplicado em nível de planejamento agrícola por Jean-Marc Boussard e Michel Petit em 1967.

Convenciona que o agricultor assuma na escolha entre possíveis culturas aquela que lhe renderá a maior margem, sujeito a restrições de terra, trabalho e capital de giro. Contudo se focaliza na admissão de que a margem mínima seja suficiente para garantir suas despesas (Foco-Perda).

A análise mostrada acima é desenvolvida a seguir:

O propósito é maximizar a expressão:

$$\sum_i C_i \cdot X_i$$

Sujeito a restrições do modelo:

$$\sum_i B_i . X_i - STBO \leq FUND \quad (1)$$

$$\sum_i C_i . X_i - LOSS - STBO = MINI \quad (2)$$

$$P_i . X_i - \frac{1}{k} LOSS \leq 0 \quad (3)$$

Em que:

- X_i Representa a i-ésima atividade de cultivo;
- C_i Representa a i-ésima unidade de margem bruta;
- B_i Representa a i-ésima unidade de capital de giro para a requerida a i-ésima atividade de cultivo;
- $FUND$ Representa a quantidade de capital de giro que o agricultor possui;
- $STBO$ Matriz que representa o valor do empréstimo realizado em curto prazo para cada atividade;
- $MINI$ Matriz que representa o mínimo de renda permitido para cada atividade;
- $LOSS$ Matriz que representa o máximo de perda admissível para cada atividade;
- $\frac{1}{k}$ Fração de perda total máxima admissível, onde o valor atribuído a k é arbitrário;
- P_i Representa a variação de risco entre culturas, diferença entre a margem bruta provável e a pior margem bruta atribuída a cada atividade agrícola, e que nunca seja superior à fração $\frac{1}{k}$.

Desta forma as equações (1) e (2), juntamente com as definições de P_i e $\frac{1}{k}$ garantem a obtenção de um resultado que, além de maximizar a função objetivo, apresenta a menor perda total.

Autores ressaltam o fato de sua aplicabilidade restringir-se a situações bastante específicas e de não se tratar de um método que se baseia em alguma medida de utilidade (DIAS, 1996; MOSS, 2010).

Mattuella (1990) enfatiza que os resultados obtidos são válidos para as propriedades pesquisadas e representadas por uma situação típica, não havendo garantias de que os mesmos se repetiriam sob outras condições.

Para Dillon (1975) esta abordagem pode ser criticada primeiramente pelo fato da escolha de k na fração de perda máxima admissível ser bastante

arbitrária, outro fato é a pressuposição de que as rendas líquidas não somente são distribuídas de modo regular mas também de forma independente.

2.3.6 PROGRAMAÇÃO LINEAR E SIMULAÇÃO MULTIDIMENSIONAL

Método desenvolvido por Dias (1996) baseado no uso de técnicas de simulação e geração de valores amostrais para um conjunto de coeficientes aleatórios que desempenham um papel relevante no problema de programação linear.

É necessário a obtenção de distribuições de probabilidade que se ajustem aos dados disponíveis. Para obter o modelo probabilístico mais adequado é fundamental uma análise prévia da quantidade e qualidade dos dados. A aplicação de técnicas de simulação em modelo de programação linear consiste na sucessiva resolução de um problema de determinístico, onde, a cada resolução, os coeficientes aleatórios são substituídos por valores obtidos através do método de geração de valores amostrais com base nas respectivas distribuições de probabilidade e possíveis correlações existentes (DIAS, 1996; FAN et al, 2001; LAW, 2007).

Depois de realizadas simulações, a análise de risco é realizada pela observação da variabilidade da distribuição da variável margem bruta total e a solução de um plano ótimo será orientada através da análise de parâmetros como assimetria, média, moda, mediana, variância e probabilidade de ocorrência de valores de margem bruta total superiores a determinados níveis mínimos estabelecidos em cada plano (DIAS, 1996; RAYCHAUDHURI, 2008; SANT'ANNA, 2011).

2.3.6.1 PROGRAMAÇÃO LINEAR

A programação linear consiste na otimização de uma função linear, denominada *função objetivo*, respeitando-se um sistema linear de igualdades ou desigualdades que recebem o nome de *restrições do modelo*.

Restrição é aquilo que impede um melhor desempenho de um sistema e representa normalmente limitações de recursos disponíveis (capital, mão-de-obra, recursos minerais ou fatores de produção) ou exigências e condições que devem ser cumpridas no problema. As restrições delimitam uma região de um plano na qual se insere o conjunto das soluções viáveis. A melhor dessas soluções, ou seja, aquela que otimiza a *função objetivo* é chamada *solução ótima*.

Segundo Caixeta-Filho (2004), a programação linear é um aprimoramento de uma técnica de resolução de sistemas de equações lineares via inversões sucessivas de matrizes com a vantagem de incorporar uma

equação linear adicional representativa de um dado comportamento que deve ser otimizada.

Dada uma função objetivo Z :

$$Z = c_1x_1 + c_2x_2 + \dots + c_dx_d$$

Sujeito às restrições:

$$a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,d}x_d \leq b_1$$

$$a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,d}x_d \geq b_2$$

...

$$a_{n,1}x_1 + a_{n,2}x_2 + \dots + a_{n,d}x_d = b_n$$

$$x_d \geq 0 \quad \text{com } d \geq 1$$

Onde a , b e c representam os parâmetros, x representa as d variáveis de decisão, sujeitas a n restrições.

Conforme explícito na Figura 4, é importante ressaltar que o mérito dado a solução do problema de programação linear deve ser associado à modelagem do problema, pois a solução apresentada será dada de acordo com a função objetivo e as restrições impostas.

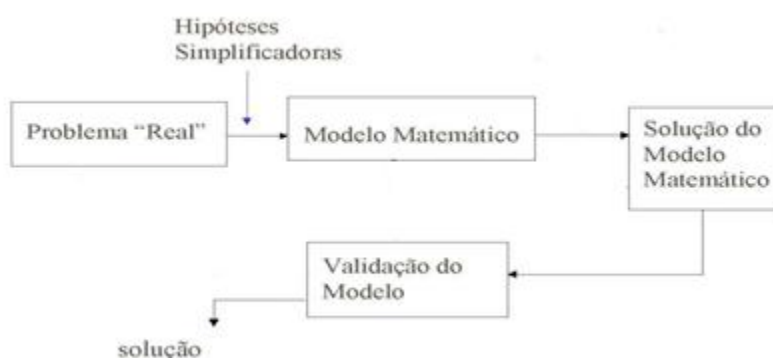


Figura 4 - Diagrama do modelo de Programação Linear

Apesar de ter uma aplicação em diversos segmentos, a programação linear não é adequada no planejamento agrícola por se tratar de um método determinístico que não considera o risco associado ao problema (CRUZ, 1984; DILLON, 1975; MOSS, 2010, PÉRES, 1981; SANT'ANNA, 2011).

2.3.6.2 OPERACIONALIZAÇÃO DO MODELO

Séries suficientemente longas seriam importantes para ajustar distribuições de probabilidade convenientes tais como, Normal, Beta, Poisson e etc, assim como estudar eventual existência de correlações entre os coeficientes aleatórios (DIAS, 1996; FAN et al, 2001, MARQUES, 2009; RAYCHAUDHURI,2008).

Da a função objetivo $Z = c_1x_1 + c_2x_2 + \dots + c_dx_d$,

e mantendo as restrições fixas, precisaremos de uma série histórica de dados que representem os coeficientes técnicos c_i com $i = 1, 2, 3, \dots, d$. Sendo representados os coeficientes técnicos como, por exemplo:

$$C_1 \Rightarrow C_{11}, C_{12}, C_{13}, \dots, C_{1n}$$

Desta forma é verificado quais distribuições de probabilidade melhor se ajustam à estas séries históricas, assim como também verificaremos possíveis correlações existentes entre as variáveis x_n e x_{n+1} (com $1 \leq n < d$) pela análise da série histórica dos parâmetros c_n e c_{n+1} (com $1 \leq n < d$) para que a partir delas, sejam gerados valores amostrais que serão incluídos no processo de simulação, resultando em m valores ótimos obtidos:

$$Z \Rightarrow Z_1, Z_2, Z_3, \dots, Z_m,$$

A análise de risco dar-se-á pela variabilidade da distribuição dos valores ótimos obtidos da variável margem bruta total (Z 's), por variabilidade entende-se a dispersão dos valores observados em torno de um valor médio.

3 MATERIAIS E MÉTODOS

No presente capítulo, mostramos uma relação existente entre os modelos de tomada de decisão no planejamento agrícola e o mercado financeiro. Em seguida trata-se da especificação dos dados, procedimentos e a estruturação do modelo. Faz-se uma descrição dos elementos da função objetivo, das restrições e dos fatores.

3.1 A TEORIA DE SELEÇÃO DE PORTFÓLIO E O MERCADO FINANCEIRO

A maioria dos artigos referentes a planejamento agrícola utiliza o modelo proposto por Markowitz em 1952 e 1959 como base na seleção de um portfólio ótimo. Entendendo-se por portfólio um conjunto de opções escolhidas entre diversas outras disponíveis em uma carteira de investimentos sujeitas a restrições, temos um problema semelhante ao de planejamento agrícola no que diz respeito à alocação de recursos escassos, mediante a restrições (DIAS,1996; MARQUES, 2009; MOSS,2010).

Segundo Choy (2011) o primeiro modelo matemático para a seleção de carteiras foi formulado por Markowitz em 1952 no seu artigo intitulado *Portfolio Selection*. Mostrando sua importância, diversos artigos da área do mercado financeiro consideram o modelo de seleção de portfólio proposto por Markowitz em 1952 como o precursor da moderna teoria de finanças (DASKALAKI et al., 2011; DRIESSEN et al., 2007; WALLS, 2004).

Segundo Markowitz (1952), uma boa carteira de investimentos é mais que um conjunto de bons ativos, é uma carteira equilibrada que permita ao investidor proteger-se e aproveitar as oportunidades mediante as inúmeras adversidades existentes no mercado financeiro. Assim, o que se pretende é encontrar a carteira mais ajustada aos objetivos do investidor.

Há dois aspectos fundamentais na análise das carteiras de investimento: a incerteza acerca da evolução futura dos ativos e a correlação existente entre a rentabilidade dos mesmos. A concentração de investimentos em um único ativo proporciona um maior risco para o investidor, enquanto uma carteira de investimentos composta por vários ativos pouco correlacionados proporciona uma redução de risco na carteira (MARKOWITZ, 1952).

3.2 CONJUNTO DE DADOS

Para a elaboração deste trabalho foi utilizado uma massa de dados referentes ao preço médio mensal *commodities* agrícolas registradas no *The European Central Bank* no período de janeiro de 1960 a setembro de 2011.

O *European Central Bank* é o banco central responsável pela moeda única europeia: o euro. Sua principal missão é preservar o poder de compra e, deste modo, a estabilidade de preços na área do euro. A área do euro é constituída pelos 17 países da União Europeia que adotaram o euro desde 1999.

Com o propósito de redução de risco na carteira, seleção das *commodities agrícolas* para a composição do portfólio foi dada a partir do coeficiente de correlação existente entre as *commodities*. Das 44 *commodities* agrícolas estudadas foram selecionadas 7 *commodities* por serem pouco ou negativamente correlacionadas.

De acordo com Markowitz (1959), opções pouco ou negativamente correlacionadas são necessárias para reduzir o risco na composição do portfólio. As *commodities* selecionadas foram as seguintes:

- *Café ICO: Preço indicativo composto - Comercializado pela ICO (International Coffee Organization);*
- *Couro Bovino - Comercializado nos Estados Unidos;*
- *Óleo Algodão - Comercializado nos Estados Unidos;*
- *Grãos de Cacau - Comercializado com preços médios diários por Nova York e Londres;*
- *Chá - Comercializado no Quênia;*
- *Juta - Comercializado em Bangladesh;*
- *Madeira Serrada - Comercializada na Malásia.*

3.3 METODOLOGIA PROPOSTA

A tecnologia agrícola proposta para aplicação no mercado de *commodities* agrícolas é a de Programação Linear e Simulação Multidimensional, que além utilizar técnicas de Programação Linear e Simulação de Monte Carlo, pode ser associada aos fundamentos teóricos da teoria de seleção de portfólio de Markowitz, introduzindo desta forma o risco no modelo de programação linear, otimizando o retorno e minimizando a variância do processo.

Neste trabalho foi utilizado o software SAS (Statistical Analysis System), no procedimento LP (Proc LP) – Linear Programming do módulo OR (Operations Research) para solucionar os modelos de programação linear, assim como o módulo básico (BASE) para gerar os números aleatórios e o módulo IML (Interactive Matrix Language) para gerar as variáveis aleatórias multinormais. O programa utilizado encontra-se no Apêndice A para variáveis correlacionadas e no Apêndice B para variáveis independentes.

3.3.1 SIMULAÇÃO DE MONTE CARLO

Segundo Raychaudhuri (2008) o método de Simulação de Monte Carlo consiste na geração de números aleatórios a partir das distribuições que se ajustam bem aos dados e, conseqüentemente, realiza-se a análise dos resultados.

São escolhidas as variáveis de estado do modelo, aquelas para descrever o problema (LAW, 2007), neste trabalho foram escolhidas sete *commodities* agrícolas, sendo utilizado como critério uma combinação entre *commodities* pouco ou negativamente correlacionadas.

No processo de simulação de Monte Carlo, desenvolve-se primeiramente um modelo determinístico, posteriormente geram-se números aleatórios a partir das distribuições de probabilidade que se ajustam bem aos dados das variáveis de entrada (RAYCHAUDHURI, 2008).

No Método de Programação Linear e Simulação Multidimensional os números gerados são correlacionados a partir das correlações existentes dos dados históricos do modelo, o conjunto formado pela geração dos valores aleatórios para cada coeficiente, a partir de suas respectivas distribuições de probabilidade, será incluído no problema de programação linear (DIAS, 1996).

A solução deste problema apontará a estratégia ótima que deverá ser seguida. Repetindo-se o processo de simulação um número suficientemente grande de vezes, os resultados observados do processo de programação linear irão compor um conjunto finito de estratégias.

O nível de confiabilidade dos resultados dependerá da representação da perfeita aleatoriedade dos coeficientes do problema de programação linear, pois as estratégias ótimas serão resultantes do efeito das distribuições individuais de probabilidade e dos coeficientes de correlação adotados na geração de números aleatórios (DIAS 1996).

Segundo Raychaudhuri (2008) o método de Monte Carlo pode ser usado na evolução de um portfólio, onde para cada simulação o comportamento dos fatores de impacto é simulado ao longo do tempo, os valores são calculados e observados a partir das características estatísticas do portfólio.

3.3.1.1 IDENTIFICAÇÃO DAS DISTRIBUIÇÕES DAS VARIÁVEIS

Para que se possa realizar o processo de simulação de Monte Carlo das variáveis aleatórias, faz-se necessário a especificação de suas distribuições de probabilidade, estas distribuições descrevem a probabilidade de ocorrência das variáveis (RAYCHAUDHURI, 2008).

Para a especificação das distribuições de probabilidade, primeiramente faz-se uma comparação entre os histogramas das distribuições teóricas e os histogramas das variáveis em análise (LAW, 2007).

Além do desenho de histogramas, existentes testes empíricos que auxiliam na escolha das distribuições de probabilidade.

No Quadro 1 é apresentado os resultados dos testes de normalidade para as sete variáveis em análise.

Quadro 1 – Teste de Normalidade para todas as variáveis

Variável – Café	Variável - Couro Bovino																																				
<table border="1"> <thead> <tr> <th colspan="3">Tests for Normality</th> </tr> <tr> <th>Test Statistic</th> <th>Value</th> <th>p-value</th> </tr> </thead> <tbody> <tr> <td>Shapiro-Wilk</td> <td>0.958751</td> <td>0.0000</td> </tr> <tr> <td>Kolmogorov-Smirnov</td> <td>0.069680</td> <td><.0100</td> </tr> <tr> <td>Cramer-von Mises</td> <td>0.420653</td> <td><.0050</td> </tr> <tr> <td>Anderson-Darling</td> <td>3.514882</td> <td><.0050</td> </tr> </tbody> </table>	Tests for Normality			Test Statistic	Value	p-value	Shapiro-Wilk	0.958751	0.0000	Kolmogorov-Smirnov	0.069680	<.0100	Cramer-von Mises	0.420653	<.0050	Anderson-Darling	3.514882	<.0050	<table border="1"> <thead> <tr> <th colspan="3">Tests for Normality</th> </tr> <tr> <th>Test Statistic</th> <th>Value</th> <th>p-value</th> </tr> </thead> <tbody> <tr> <td>Shapiro-Wilk</td> <td>0.911988</td> <td>0.0000</td> </tr> <tr> <td>Kolmogorov-Smirnov</td> <td>0.137902</td> <td><.0100</td> </tr> <tr> <td>Cramer-von Mises</td> <td>2.951401</td> <td><.0050</td> </tr> <tr> <td>Anderson-Darling</td> <td>18.86479</td> <td><.0050</td> </tr> </tbody> </table>	Tests for Normality			Test Statistic	Value	p-value	Shapiro-Wilk	0.911988	0.0000	Kolmogorov-Smirnov	0.137902	<.0100	Cramer-von Mises	2.951401	<.0050	Anderson-Darling	18.86479	<.0050
Tests for Normality																																					
Test Statistic	Value	p-value																																			
Shapiro-Wilk	0.958751	0.0000																																			
Kolmogorov-Smirnov	0.069680	<.0100																																			
Cramer-von Mises	0.420653	<.0050																																			
Anderson-Darling	3.514882	<.0050																																			
Tests for Normality																																					
Test Statistic	Value	p-value																																			
Shapiro-Wilk	0.911988	0.0000																																			
Kolmogorov-Smirnov	0.137902	<.0100																																			
Cramer-von Mises	2.951401	<.0050																																			
Anderson-Darling	18.86479	<.0050																																			
Variável – Óleo Algodão	Variável – Grão de Cacau																																				
<table border="1"> <thead> <tr> <th colspan="3">Tests for Normality</th> </tr> <tr> <th>Test Statistic</th> <th>Value</th> <th>p-value</th> </tr> </thead> <tbody> <tr> <td>Shapiro-Wilk</td> <td>0.894705</td> <td>0.0000</td> </tr> <tr> <td>Kolmogorov-Smirnov</td> <td>0.094325</td> <td><.0100</td> </tr> <tr> <td>Cramer-von Mises</td> <td>1.215935</td> <td><.0050</td> </tr> <tr> <td>Anderson-Darling</td> <td>9.514673</td> <td><.0050</td> </tr> </tbody> </table>	Tests for Normality			Test Statistic	Value	p-value	Shapiro-Wilk	0.894705	0.0000	Kolmogorov-Smirnov	0.094325	<.0100	Cramer-von Mises	1.215935	<.0050	Anderson-Darling	9.514673	<.0050	<table border="1"> <thead> <tr> <th colspan="3">Tests for Normality</th> </tr> <tr> <th>Test Statistic</th> <th>Value</th> <th>p-value</th> </tr> </thead> <tbody> <tr> <td>Shapiro-Wilk</td> <td>0.920606</td> <td>0.0000</td> </tr> <tr> <td>Kolmogorov-Smirnov</td> <td>0.136866</td> <td><.0100</td> </tr> <tr> <td>Cramer-von Mises</td> <td>2.246369</td> <td><.0050</td> </tr> <tr> <td>Anderson-Darling</td> <td>12.45462</td> <td><.0050</td> </tr> </tbody> </table>	Tests for Normality			Test Statistic	Value	p-value	Shapiro-Wilk	0.920606	0.0000	Kolmogorov-Smirnov	0.136866	<.0100	Cramer-von Mises	2.246369	<.0050	Anderson-Darling	12.45462	<.0050
Tests for Normality																																					
Test Statistic	Value	p-value																																			
Shapiro-Wilk	0.894705	0.0000																																			
Kolmogorov-Smirnov	0.094325	<.0100																																			
Cramer-von Mises	1.215935	<.0050																																			
Anderson-Darling	9.514673	<.0050																																			
Tests for Normality																																					
Test Statistic	Value	p-value																																			
Shapiro-Wilk	0.920606	0.0000																																			
Kolmogorov-Smirnov	0.136866	<.0100																																			
Cramer-von Mises	2.246369	<.0050																																			
Anderson-Darling	12.45462	<.0050																																			

Variável – Chá	Variável – Juta																																				
<table border="1"> <thead> <tr> <th colspan="3">Tests for Normality</th> </tr> <tr> <th>Test Statistic</th> <th>Value</th> <th>p-value</th> </tr> </thead> <tbody> <tr> <td>Shapiro-Wilk</td> <td>0.897078</td> <td>0.0000</td> </tr> <tr> <td>Kolmogorov-Smirnov</td> <td>0.115517</td> <td><.0100</td> </tr> <tr> <td>Cramer-von Mises</td> <td>1.078435</td> <td><.0050</td> </tr> <tr> <td>Anderson-Darling</td> <td>7.049093</td> <td><.0050</td> </tr> </tbody> </table>	Tests for Normality			Test Statistic	Value	p-value	Shapiro-Wilk	0.897078	0.0000	Kolmogorov-Smirnov	0.115517	<.0100	Cramer-von Mises	1.078435	<.0050	Anderson-Darling	7.049093	<.0050	<table border="1"> <thead> <tr> <th colspan="3">Tests for Normality</th> </tr> <tr> <th>Test Statistic</th> <th>Value</th> <th>p-value</th> </tr> </thead> <tbody> <tr> <td>Shapiro-Wilk</td> <td>0.718729</td> <td>0.0000</td> </tr> <tr> <td>Kolmogorov-Smirnov</td> <td>0.198429</td> <td><.0100</td> </tr> <tr> <td>Cramer-von Mises</td> <td>8.764947</td> <td><.0050</td> </tr> <tr> <td>Anderson-Darling</td> <td>49.24486</td> <td><.0050</td> </tr> </tbody> </table>	Tests for Normality			Test Statistic	Value	p-value	Shapiro-Wilk	0.718729	0.0000	Kolmogorov-Smirnov	0.198429	<.0100	Cramer-von Mises	8.764947	<.0050	Anderson-Darling	49.24486	<.0050
Tests for Normality																																					
Test Statistic	Value	p-value																																			
Shapiro-Wilk	0.897078	0.0000																																			
Kolmogorov-Smirnov	0.115517	<.0100																																			
Cramer-von Mises	1.078435	<.0050																																			
Anderson-Darling	7.049093	<.0050																																			
Tests for Normality																																					
Test Statistic	Value	p-value																																			
Shapiro-Wilk	0.718729	0.0000																																			
Kolmogorov-Smirnov	0.198429	<.0100																																			
Cramer-von Mises	8.764947	<.0050																																			
Anderson-Darling	49.24486	<.0050																																			
Variável – Madeira Serrada																																					
<table border="1"> <thead> <tr> <th colspan="3">Tests for Normality</th> </tr> <tr> <th>Test Statistic</th> <th>Value</th> <th>p-value</th> </tr> </thead> <tbody> <tr> <td>Shapiro-Wilk</td> <td>0.945360</td> <td>0.0000</td> </tr> <tr> <td>Kolmogorov-Smirnov</td> <td>0.119171</td> <td><.0100</td> </tr> <tr> <td>Cramer-von Mises</td> <td>1.676256</td> <td><.0050</td> </tr> <tr> <td>Anderson-Darling</td> <td>9.636931</td> <td><.0050</td> </tr> </tbody> </table>	Tests for Normality			Test Statistic	Value	p-value	Shapiro-Wilk	0.945360	0.0000	Kolmogorov-Smirnov	0.119171	<.0100	Cramer-von Mises	1.676256	<.0050	Anderson-Darling	9.636931	<.0050																			
Tests for Normality																																					
Test Statistic	Value	p-value																																			
Shapiro-Wilk	0.945360	0.0000																																			
Kolmogorov-Smirnov	0.119171	<.0100																																			
Cramer-von Mises	1.676256	<.0050																																			
Anderson-Darling	9.636931	<.0050																																			

O teste apresentou-se significativo para as sete variáveis em questão indicando que nenhuma das distribuições de dados segue a distribuição normal.

Segundo Dias (1996) a distribuição triangular é sugerida devido a sua capacidade em expressar distribuições com assimetria, entretanto é necessário saber o valor de máximo, de mínimo e o valor modal.

Neste trabalho foi utilizado o modelo triangular para ajuste da distribuição de probabilidade dos coeficientes estocásticos.

Segundo Law (2007), sendo como parâmetros o valor de mínimo (a), máximo (b) e a moda (m), a função triangular possui a seguinte função densidade:

$$f(x) = \begin{cases} \frac{2(X-a)}{(m-a)(b-a)} & (\text{Se } a \leq X \leq m) \\ \frac{2(b-X)}{(b-m)(b-a)} & (\text{Se } m \leq X \leq b) \\ 0 & (\text{Se } X < a \text{ ou } X > b) \end{cases}$$

A Figura 5 mostra o gráfico da função densidade da distribuição triangular:

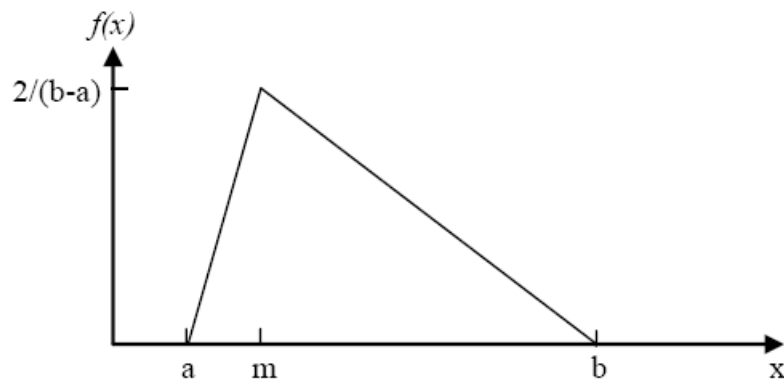


Figura 5 - Gráfico da densidade triangular Triang(a, b, m)

Identificadas às distribuições de probabilidade que serão usadas no modelo, o processo de simulação terá por objetivo gerar valores para os coeficientes aleatórios (FAN *et al.* 2001; LAW, 2007).

3.3.1.2 GERAÇÃO DE NÚMEROS ALEATÓRIOS

As simulações foram realizadas com de variáveis independentes e variáveis correlacionadas, neste último caso necessário o uso de uma matriz de correlações para a geração das variáveis (DIAS, 1996; DIAS *et al.*, 2008).

O método consiste em gerar variáveis aleatórias correlacionadas com distribuição normal, após isto segue para a geração de variáveis distribuídas uniformemente entre (0,1), para depois transformá-las em valores com as distribuições de interesse (DIAS *et al.*, 2008), neste caso a distribuição triangular.

Dias (1996), Dias *et al.* (2008) e Sant'Anna (2011) fazem uma boa revisão do processo de geração de números aleatórios, citando:

3.3.1.2.1 GERAÇÃO DE NÚMEROS ALEATÓRIOS COM DISTRIBUIÇÃO NORMAL

Por se tratar de variáveis correlacionadas, utilizou-se o método das distribuições condicionais.

Uma distribuição k-dimensional equivale ao produto de k distribuições condicionais que, através da geração em sequência de observações de distribuições condicionais univariadas, obtêm-se variáveis aleatórias com distribuição multivariada (DIAS, 1996).

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

Segundo Dias *et al.*(2008), se X segue uma distribuição normal multivariada não singular com vetor de médias μ temos que:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} & \dots & x_{1n} \\ x_{21} & x_{21} & x_{21} & \dots & \dots & x_{2n} \\ x_{31} & x_{31} & \dots & \dots & \dots & x_{3n} \\ x_{41} & \dots & \cdot & \cdot & \cdot & x_{31} \\ \dots & \cdot & \cdot & \cdot & \cdot & \dots \\ x_{k1} & \cdot & \cdot & \cdot & \cdot & x_{kn} \end{bmatrix} = (x_1, x_2, x_3, \dots, x_k)'$$

Com:

$$\bullet \mu = (\mu_1, \mu_2, \mu_3, \dots, \mu_k)'$$
 (2)

E a matriz de dispersão semidefinida positiva:

$$\bullet \Sigma = \sigma_{ij}$$
 (3)

Com o i-ésimo elemento da diagonal Σ sendo:

$$\bullet \sigma_i^2 = \sigma_{ii}$$
 (4)

Para $i = 1, 2, 3, \dots, k$, definimos:

- $X_i = (x_1, x_2, x_3, \dots, x_k)'$

Como o vetor dos primeiros i 's componentes de X , e seu vetor de médias é dado por:

- $\mu_i = (\mu_1, \mu_2, \mu_3, \dots, \mu_i)'$ (6)

A matriz de dispersão de X_i é dada por:

- $\Sigma_{ii} = \begin{bmatrix} \sigma_{11} & \dots & \sigma_{1i} \\ \vdots & \ddots & \vdots \\ \sigma_{i1} & \dots & \sigma_{ii} \end{bmatrix} = \begin{bmatrix} \text{var}(x_1) & \dots & \text{cov}(x_1, x_i) \\ \vdots & \ddots & \vdots \\ \text{cov}(x_i, x_1) & \dots & \text{var}(x_i) \end{bmatrix}$ (7)

É a submatriz i por i de Σ .

Sendo assim, o vetor de covariância X_{i+1} com $X_1, X_2, X_3, \dots, X_i$ é dado por:

- $\sigma_{(i)} = (\sigma_{1,i+1}, \sigma_{2,i+1}, \sigma_{3,i+1}, \dots, \sigma_{i,i+1})'$ (8)

Segundo Dias (1996), a distribuição marginal de X_1 é normal univariada com média μ_1 e variância σ_1^2 , ao passo que i vai de 1 até $k - 1$, a distribuição condicional de X_{i+1} , dada $X_1, X_2, X_3, \dots, X_i$, é normal e univariada com média $\bar{\mu}_{i+1}$ e variância $\bar{\sigma}_{(i+1)}^2$ dadas por:

- $\bar{\mu}_{i+1} = \mu_{i+1} + \sigma'_{(i)} \Sigma_{ii}^{-1} (x_{(i)} - \mu_{(i)})$ (9)

- $\bar{\sigma}_{(i+1)}^2 = \sigma_{(i+1)}^2 - \sigma'_{(i)} \Sigma_{ii}^{-1} \sigma_{(i)}$ (10)

Desta forma, temos o seguinte algoritmo para gerar números aleatórios com distribuição normal multivariada k -dimensional com média μ e matriz de dispersão Σ (DIAS, 1996).

- Passo 1. Faça $i = 1$, $\mu = \mu_1$ e $\sigma^2 = \sigma_1^2$
- Passo 2. Gere $X_i = x_1$ de uma $N(\mu, \sigma^2)$
- Passo 3. Se $i = k$, então atribua $\mathbf{X} = (x_1, x_2, x_3, \dots, x_k)'$
- Passo 4. Incremente i
- Passo 5. Calcule $\mu = \bar{\mu}_i$ e $\sigma^2 = \bar{\sigma}_i^2$
- Passo 6. Volte ao passo 2.

3.3.1.2.2 GERAÇÃO DE NÚMEROS ALEATÓRIOS COM DISTRIBUIÇÃO UNIFORME

Segundo Dias *et al.*(2008), sendo $(x_1, x_2, x_3, \dots, x_p)$ uma amostra aleatória de uma distribuição normal multivariada com a seguinte matriz de covariância:

$$V = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} & \dots & \rho_{1p} \\ \rho_{21} & 1 & \rho_{23} & \dots & \dots & \rho_{2p} \\ \rho_{31} & \rho_{32} & 1 & \dots & \dots & \dots \\ \dots & \dots & \dots & 1 & \dots & \dots \\ \dots & \dots & \dots & \dots & 1 & \dots \\ \rho_{p1} & \rho_{p2} & \dots & \dots & \dots & 1 \end{bmatrix} \quad \text{e matriz de médias } \mu = \begin{bmatrix} 0 \\ 0 \\ \dots \\ \dots \\ \dots \\ 0 \end{bmatrix},$$

Neste caso os valores gerados com distribuição uniforme (u) serão obtidos a partir da geração de valores com distribuição normal.

Se $F(x)$ é a distribuição acumulada de uma variável normal padrão, temos $U_i = F(x_i)$ como uma distribuição uniforme no intervalo (0,1) com probabilidade de valores menores ou iguais a X_i e distribuídos uniformemente para todo i (DIAS *et al.*, 2008), com correlação entre U_i e U_j igual a :

$$\bullet \rho_{U_{ij}} = \left(\frac{6}{\pi}\right) \cdot \arcsen\left(\frac{\rho_{ij}}{2}\right) \quad (1)$$

No intervalo (0,1) ρ_{ij} e $\rho_{U_{ij}}$ são similares, sendo que a maior diferença entre os dois é aproximadamente 0,02. Sendo assim, podemos considerar que a correlação entre x_i e x_j seja igual a $\rho_{U_{ij}}$, assim da expressão acima temos:

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

Utilizaremos o método da transformação-inversa para gerar as variáveis uniformes correlacionadas, sendo assim a geração de uma amostra aleatória X contínua, com função de distribuição F contínua e estritamente crescente em $0 < F(x) < 1$, cuja inversa é $F^{-1}(x)$, será dada pela seguinte sequência (SANT'ANNA, 2011):

Passo 1. Gere $U \sim U(0,1)$

Passo 2. Retorne $X = F^{-1}(U)$

Como $F^{-1}(U)$ é definida, pois $0 \leq U \leq 1$ e o domínio de F é $[-1, 1]$, os valores aleatórios gerados assumem valores positivos ou negativos dependendo do valor de U conforme é apresentado na Figura 7.

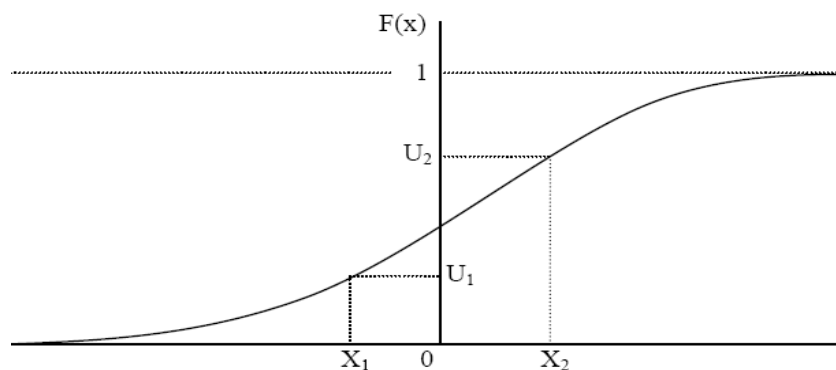


Figura 6 - Ilustração do método da transformação inversa para variáveis aleatórias contínuas

A geração de variáveis aleatórias com outras distribuições de probabilidade é feita a partir da geração de variáveis aleatórias correlacionadas com distribuição uniforme (DIAS et al., 2008; SANT'ANNA, 2011).

3.3.1.2.3 GERAÇÃO DE NÚMEROS ALEATÓRIOS COM DISTRIBUIÇÃO TRIANGULAR

Os valores com distribuição triangular são obtidos a partir de valores com distribuição uniforme no intervalo (0,1). Para isto aplica-se a função de distribuição inversa $F_T^{-1}(U) = T$. Desta forma teremos uma variável com distribuição triangular (DIAS, 1996).

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

Na Tabela 1 são apresentados os parâmetros utilizados na geração de números aleatórios que representaram as sete *commodities* agrícolas.

Tabela 1 - Parâmetros de mínimo, máximo e moda utilizados na geração de números aleatórios com distribuição triangular

Commodity	Variável	Parâmetros
Café ICO	x_1	(64,1; 95,1; 359,9)
Couro Bovino	x_2	(9,2; 82,3; 127,5)
Óleo Algodão	x_3	(40,5; 46,8; 399,7)
Grãos de Cacau	x_4	(68,9; 157,6; 552,7)
Chá	x_5	(17,2; 103,4; 153,8)
Juta	x_6	(65,3; 77,4; 150,7)
Madeira Serrada	x_7	(66; 95,8; 385,5)

Desta forma, pode-se utilizar o seguinte algoritmo de transformação inversa para gerar $X \sim \text{triang}(a, m, b)$:

Passo 1. Gere $U \sim U(0,1)$

Passo 2. Se: $0 \leq U \leq \frac{m-a}{b-a}$, então $X = \sqrt{mu}$, caso contrário,
 $X = b - \sqrt{(1-u)(b-a)(a-b)}$

Assim foram gerados 1000 números aleatórios para cada uma das sete variáveis em questão: *commodities* agrícolas.

3.3.1.2.4 FORMULAÇÃO MATEMÁTICA PARA A GERAÇÃO DE NÚMEROS ALEATÓRIOS A PARTIR DO MODELO DE MARKOWITZ

Segundo Markowitz (1952), uma boa carteira de investimentos é uma carteira equilibrada que permita ao investidor proteger-se e aproveitar as oportunidades face a adversidades. Propõe a maximização do retorno concomitantemente com a minimização do risco, onde os dados de entrada necessários para o procedimento são o retorno médio de cada ativo, a matriz de covariância da taxa de retorno dos ativos e a restrição de positividade dos investimentos.

Algumas considerações devem ser realizadas na formulação matemática do modelo proposto por Markowitz (CHOY, 2011; MARKOWITZ, 1959):

- A correlação entre os ativos são fixas e constantes;
- O investidor procura maximizar o lucro;
- O investidor é adverso ao risco.

Para maximizar a relação retorno/risco, a formulação matemática é dada da seguinte forma:

Maximizar:

$$\sum_{i=1}^n \sum_{j=1}^n \frac{R_i x_i}{\sigma_{ij} x_i x_j}$$

Onde:

- R_i É o taxa de retorno do ativo;
- n Número de ativos;
- σ_{ij} Variância do ativo i quando $i=j$, e covariância entre os ativos i e j quando $i \neq j$;

- x_j Proporção de participação de cada ativo na composição da carteira.

A expressão citada acima para as sete *commodities* agrícolas temos:

$$\text{Máx: } \sum_{j=1}^7 \sum_{i=1}^7 \frac{R_i}{x_j \sigma_{ij}} \quad (1)$$

Sujeito a restrições de positividade de investimentos:

- $x_j \geq 0$ para $j = 1, 2, 3, \dots, n$

E restrições de investimentos:

- $\sum_{j=1}^7 x_j = 100\%$

Uma vez que a proporção de investimento encontra-se no denominador, devemos substituímos a expressão $\frac{1}{x_j}$ por y_j para a formulação no processo de programação linear, deixando desta forma o inverso da proporção de investimentos no numerador.

Desta forma, maximizar a expressão (1) é o mesmo que minimizar a expressão (2) já que x_j e y_j são inversamente proporcionais.

$$\text{Mín: } \sum_{j=1}^7 \sum_{i=1}^7 y_j \frac{R_i}{\sigma_{ij}} \quad (2)$$

Segundo Markowitz (1952) a taxa de retorno R_i é dada pela diferença entre o valor gerado R_c (número aleatório criado a partir da distribuição de probabilidade) e o valor esperado \bar{R}_l (média da série histórica de dados), o retorno esperado é dado pela soma ponderada dos retornos individuais i :

$$\bullet R_i = R_c - \bar{R}_l \quad (3)$$

$$\bullet R_e = x_1 R_1 + x_2 R_2 + x_3 R_3 + \dots + x_i R_i \quad (4)$$

E a variância (risco) do processo é dada por:

$$\bullet V = \sum_{i=1}^7 \sum_{j=1}^7 \sigma_{ij} x_i x_j \quad (5)$$

4 RESULTADOS E DISCUSSÃO

Neste capítulo são apresentados os resultados das análises efetuadas a partir do processo de geração de números aleatórios e das simulações. A discussão envolve o processo de tomada de decisão para a alocação de recursos no mercado financeiro a partir do modelo proposto: Programação Linear e Simulação Multidimensional.

4.1 RESULTADOS

Foram gerados 1000 números aleatórios com distribuição triangular para cada variável. A variável t1 representa as variáveis criadas a partir dos parâmetros da commodity agrícola Café, a t2 representa as variáveis criadas a partir dos parâmetros da commodity agrícola Couro Bovino, a t3 representa as variáveis criadas a partir dos parâmetros da commodity agrícola óleo Algodão, a t4 representa as variáveis criadas a partir dos parâmetros da commodity agrícola Grãos de Cacau, t5 representa as variáveis criadas a partir dos parâmetros da commodity agrícola Madeira Serrada, t6 representa as variáveis criadas a partir dos parâmetros da commodity agrícola chá e t7 representa as variáveis criadas a partir dos parâmetros da commodity agrícola Juta. Na Tabela 2 é apresentado o coeficiente de correlação destas variáveis.

Tabela 2 - Coeficiente de correlação de Pearson para as variáveis geradas

	t1	t2	t3	t4	t5	t6	t7
t1	1	-0.23	0.3	0.64	-0.29	0.55	0.34
t2	-0.23	1	0.45	-0.19	0.76	-0.24	0.15
t3	0.3	0.45	1	0.33	0.22	0.45	0.42
t4	0.64	-0.19	0.33	1	-0.36	0.67	0.39
t5	-0.29	0.76	0.22	-0.36	1	-0.29	0.02
t6	0.55	-0.24	0.45	0.67	-0.29	1	0.62
t7	0.34	0.15	0.42	0.39	0.02	0.62	1

O valor -0.19 significa que existe correlação negativa existente entre as *commodities* agrícolas Couro Bovino e Grãos de Cacau, já o valor 0.62 representa a correlação existente entre as *commodities* agrícolas Chá e Juta.

Na Figura 7 é apresentado o histograma da variável dependente gerada com distribuição triangular que representa a commodity Café ICO.

São observados valores com alta frequência a partir do número 80 que permanece até o valor 140 com uma diminuição nas frequências a partir do número 180. Observamos pelo Histograma dos números gerados representam uma distribuição triangular.

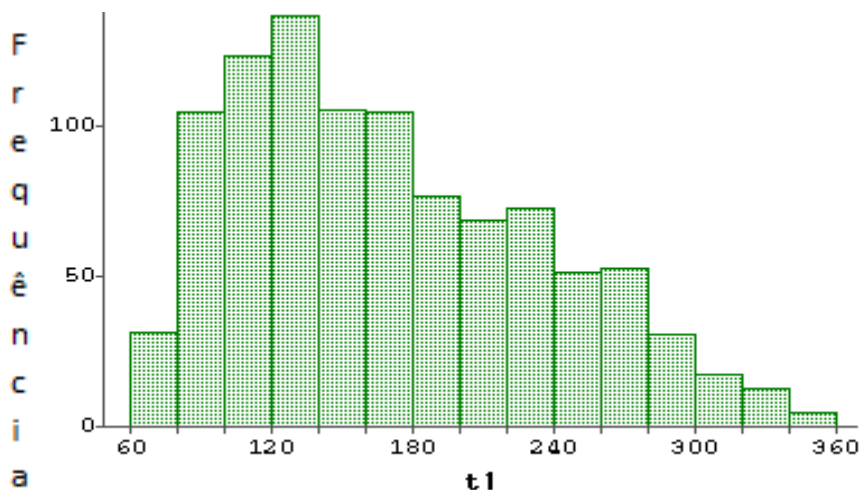


Figura 7 - Histograma da variável Café IOC gerada com distribuição triangular

Na Figura 8 é apresentado o histograma da variável dependente gerada com distribuição triangular que representa a *commoditie* Couro Bovino. Observa-se um crescimento contínuo com leve oscilação dos valores gerados até o valor 80 que tem a maior frequência onde, a partir daí, temos um forte decréscimo na frequência. O valor 128 apresenta-se como máximo e com uma baixa frequência.

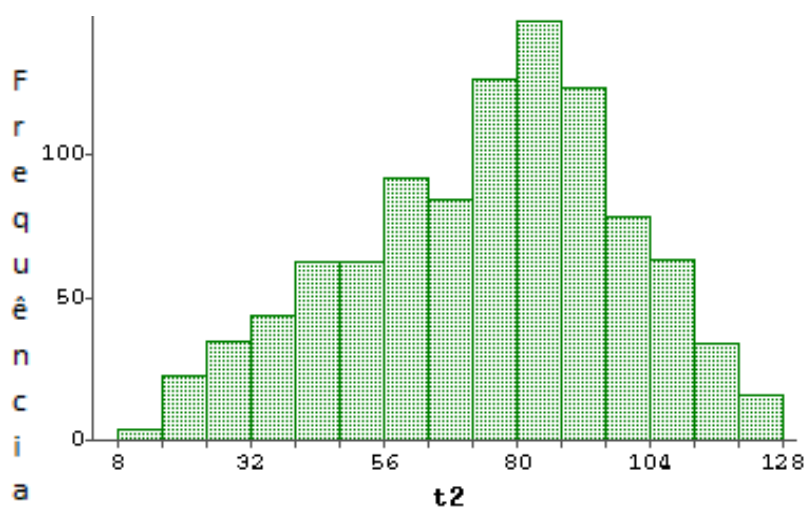


Figura 8 - Histograma da variável Couro Bovino gerada com distribuição triangular

Na Figura 9 podemos observar pelo histograma apresentado da variável dependente gerada com distribuição triangular que representa a *commoditie* Óleo Algodão um ponto de máxima frequência a partir do valor 62.5 tornando-se constante até o valor 87.5 onde observamos um decaimento na distribuição de frequências.

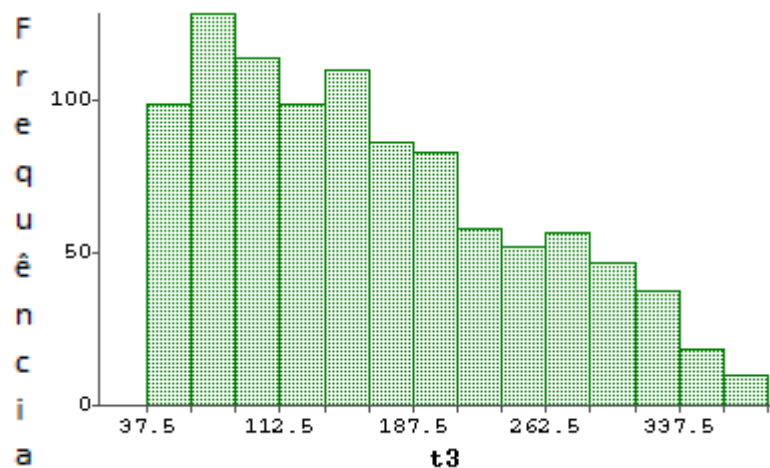


Figura 9 - Histograma da variável Óleo Algodão gerada com distribuição triangular

Na Figura 10 é apresentado o histograma que representa a *commoditie* Grãos de Cacau, Podemos observar uma alta frequência entre os valores 90 e 180 e baixa frequência com leves oscilações a partir deste valor.

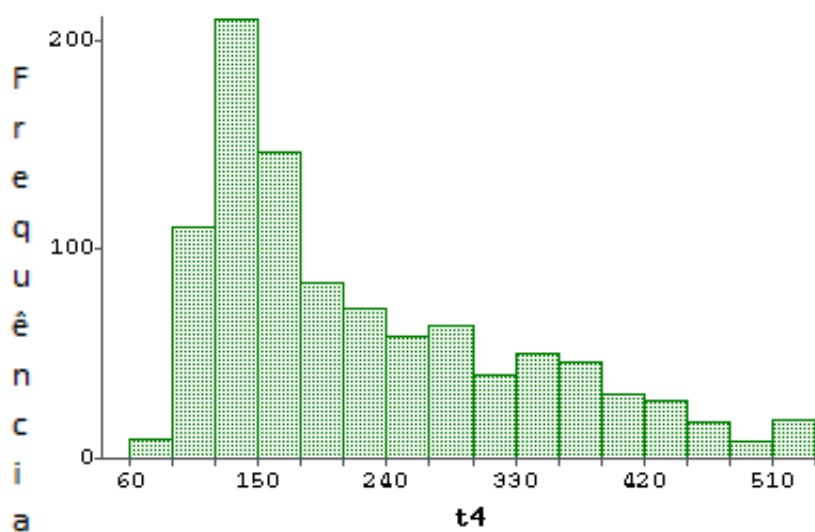


Figura 10 - Histograma da variável Grãos de Cacau gerada com distribuição triangular

Como mostra a Figura 11, o histograma apresentado da variável dependente gerada com distribuição triangular que representa a commodity Madeira Serrada, pode-se observar que os valores acima da média (média=93.37) apresentam-se com maior frequência.

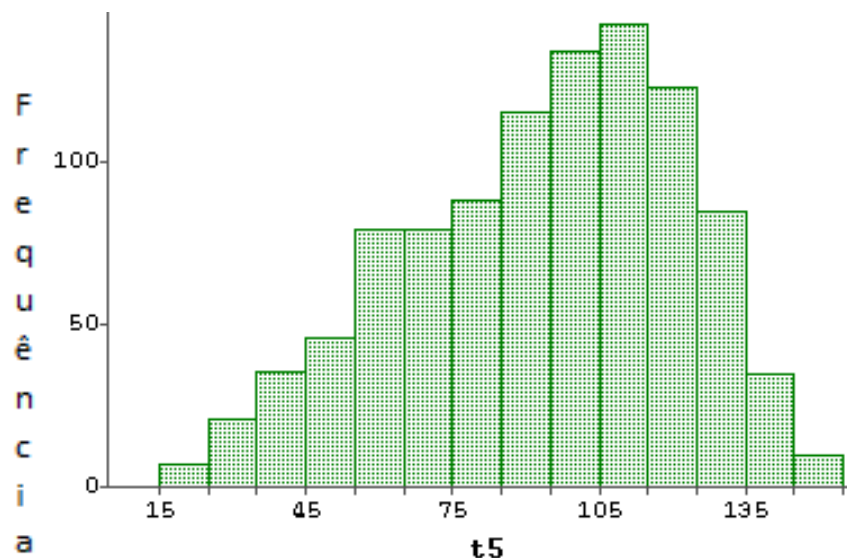


Figura 11 - Histograma da variável Madeira Serrada gerada com distribuição triangular

De acordo com a Figura 12 é apresentado o histograma da variável dependente gerada com distribuição triangular que representa a commodity agrícola Chá. A variável apresenta altas frequências em valores próximos à média (Média = 97.5) com pequenas oscilações.

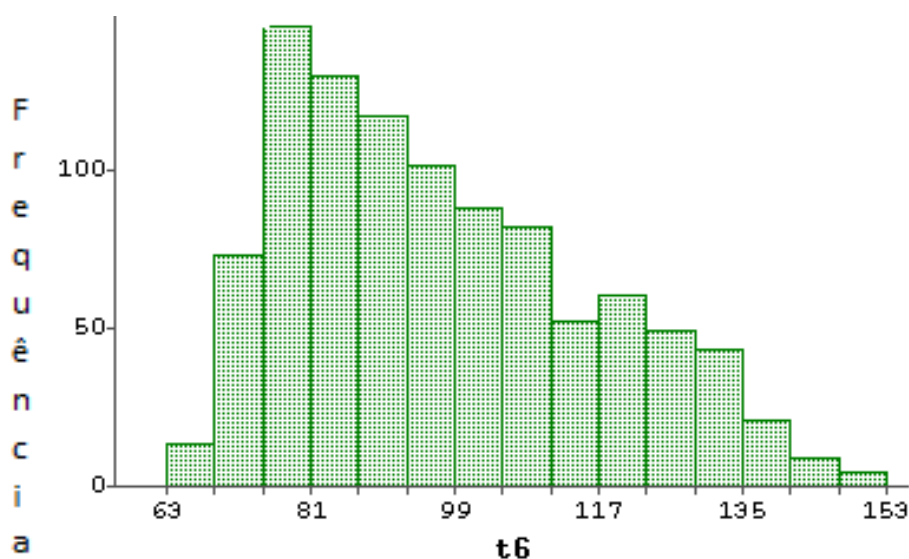


Figura 12 - Histograma da variável Chá gerada com distribuição triangular

Na Figura 13 é apresentado o histograma da variável dependente gerada com distribuição triangular que representa a *commoditie* juta. A variável apresenta valores médios com altas frequências próximos à média (Média = 185.9) com um decaimento contínuo até os altos valores.

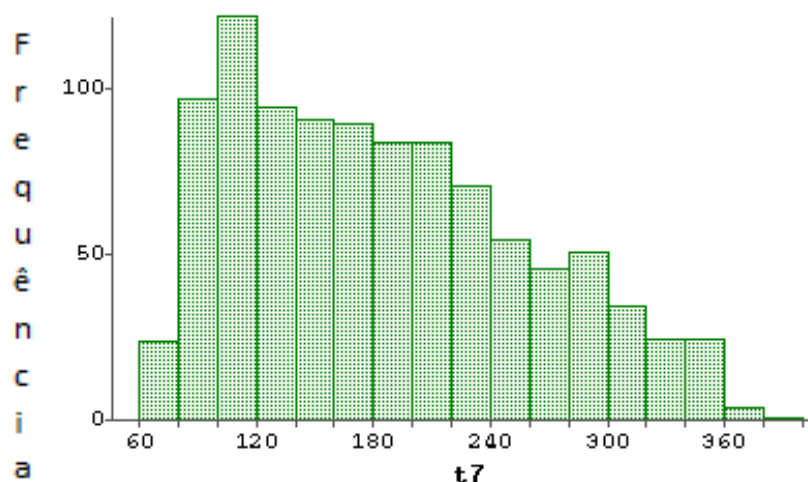


Figura 13 - Histograma da variável Juta gerada com distribuição triangular

Estatísticas como média, moda, variância e assimetria dão suporte ao processo de tomada de decisão (DIAS, 1996; RAYCHAUDHURI, 2008). Considerando a média, são apresentados na tabela 4 os resultados para os coeficientes com e sem dependência entre as variáveis, são dados em porcentagens de investimentos para cada *commoditie* agrícola.

Tabela 3 – Taxas médias de alocação investimentos nas *commodities* agrícolas

<i>Commodities</i>	<i>Café ICO</i>	<i>Couro Bovino</i>	<i>Óleo Algodão</i>	<i>Grãos de Cacau</i>	<i>Madeira Serrada</i>	<i>Chá</i>	<i>Juta</i>
Com dependência	15.30%	15.6%	0.8%	10.1%	16.8%	27.6%	13.8%
Sem dependência	8.8%	16.1%	1.4%	4.8%	18.9%	32.5%	17.5%

As estatísticas para a margem bruta mostram que esta variável apresenta assimetria negativa, confirmada pelo teste de normalidade, que mostra que a distribuição é estatisticamente diferente da normal. A partir dos

valores de mínimo e máximo, e das medidas de dispersão, observamos que se trata de um processo com risco considerável.

Na tabela 3 são apresentadas estatísticas simples para a margem bruta obtidas a partir dos resultados de simulação com e sem dependência entre os coeficientes do modelo de programação linear. As medidas de dispersão foram bastante próximas mostrando que a dependência entre as variáveis não foi determinante na análise de risco.

Tabela 4 - Estatísticas simples para a margem bruta

Estatísticas	Sem dependência	Com dependência
Média	-47.53	-39.55
Moda	-----	-----
Mediana	-39.61	-33.66
Assimetria	-1.14	-1.09
Curtose	1.86	1.70
Desvio Padrão	41.76	39.99
Variância	1744.61	1599.92
Erro Padrão	1.32	1.26
Coefficiente de variação	-87.86	-101.13
Mínimo	-251.308	-210.093
Máximo	20.81	27.34
Teste de Normalidade	Dist. Não Normal Shapiro-Wilk / W=0.92879 Pr < W <0.0001	Dist. Não Normal Shapiro-Wilk / W=0.934156 Pr < W <0.0001

Nas figuras 14 e 15 é apresentado o histograma da distribuição de frequência da margem bruta com e sem dependência entre as variáveis do modelo de programação linear.

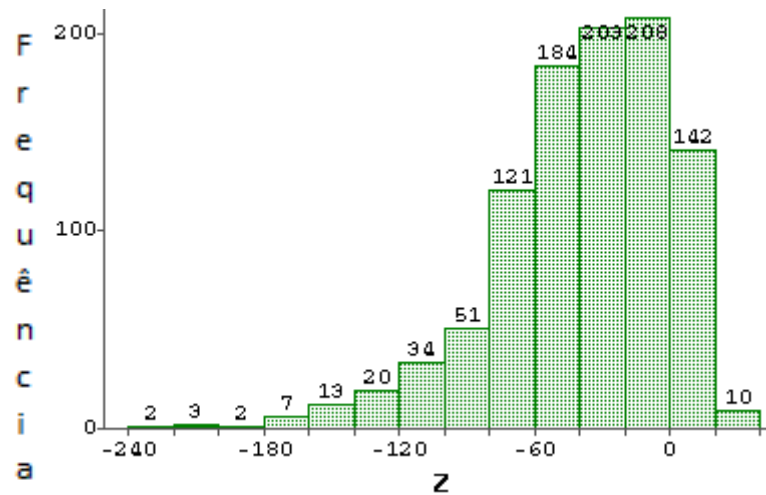


Figura 14 - Histograma da margem bruta com dependência

Observamos um comportamento levemente mais oscilante na distribuição de frequência com dependência do que sem dependência entre variáveis refletindo a diminuição na variabilidade na distribuição da margem bruta total e conseqüentemente o risco.

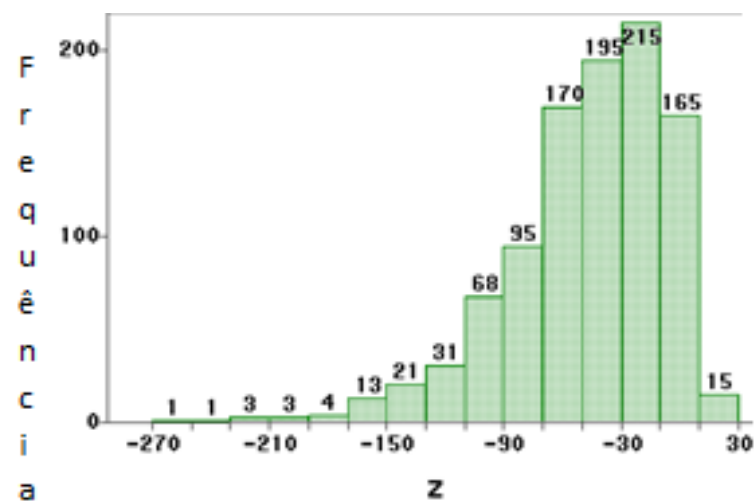


Figura 15 - Histograma da margem bruta sem dependência

4.2 CONCLUSÕES E SUGESTÕES

Este trabalho teve como objetivo principal aplicar uma metodologia de auxílio ao planejamento agrícola no mercado financeiro cujos ativos são *commodities* agrícolas. Após a composição da carteira e concluídas as simulações realizamos as seguintes conclusões:

- O uso da simulação mostrou-se consistente como instrumento de análise de risco e dos efeitos de aleatoriedade dos coeficientes estocásticos no processo de simulação;
- O processo de tomada de decisão através de técnicas de programação linear e geração de números aleatórios correlacionados aumentam o realismo dos modelos de decisão ao torna-los estocásticos;
- O uso de simulação sem dependência não subestimou o risco quando comparado com a simulação com dependência;
- O fato de que a cada simulação os resultados se mostraram diferentes, apesar de próximos, faz com que a utilização desta tecnologia agrícola no mercado financeiro mostra-se satisfatória apenas como ferramenta de suporte a decisão devendo ser utilizada em conjunto a outras ferramentas de previsão;
- O fato da formulação matemática do modelo de programação linear para mercado financeiro não possuir muitas restrições foi determinante para que a escolha entre as alternativas fosse a média, tendo em vista que a cada simulação o programa concentrava todo o investimento na *commodity* agrícola mais rentável;
- A utilização da tecnologia agrícola no mercado financeiro não possibilitou a percepção do melhor momento para realizar o investimento, ou seja, comprar e vender as *commodities* agrícolas;
- Sugere-se que, a partir da série de retorno, seja realizado um estudo sobre o comportamento das altas e baixas frequências com o objetivo de se analisar a volatilidade e a tendência do mercado para perceber o melhor momento para realizar o investimento;

- Sugere-se realizar uma análise de autocorrelação na série histórica de dados de todas as *commodities* com o propósito refletir esta autocorrelação na geração de números aleatórios;
- Sugere-se que outros modelos de tomada de decisão que incluem o risco no processo de programação linear sejam utilizados com a inclusão do processo de simulação e geração de números aleatórios correlacionados.
- Sugere-se que seja realizado um estudo com distribuições que se ajustem melhor aos dados com o propósito de se obter resultados mais próximos da realidade.

REFERÊNCIAS

ANDERSON, J.R. Risk Efficiency in the interpretation of agricultural Production Research. **Review of Marketing and Agricultural Economics**, p.131-184, September, 1974.

BITTENCOURT, M. V. L.; SAMPAIO, A. V. **Análise teórica dos modelos de tomada de decisão sob condições de risco**. CMDE/UFPR, Curitiba, 1998. (Texto para discussão 01/98)

BOUSSARD, J. M.; M. PETIT. Representation of Farmers Behaviour under Uncertainty with a Focus-Loss Constraint. **J. Farm Econ.**, n.49, v.4, p.869-80, 1967.

CAIXETA FILHO, J.V. **Pesquisa operacional: técnicas de otimização aplicadas a sistemas agroindustriais**. 2ª ed. São Paulo: Atlas, 2004.

CHOY, MURPHY; Modern Portfolio Theory using SAS. In: ESUG, Agosto 22-26, 2011, Barcelona, ESP. **Anais**, Barcelona, ESP, 2011

CUSINATO R.T. **Teoria da decisão sob incerteza e a hipótese da utilidade esperada: Conceitos analíticos e paradoxos**. Dissertação de Mestrado em Economia. UFRGS, Rio Grande do Sul, 2003. 181p.

CRUZ, E. R. da. **Aspectos teóricos sobre incorporação de riscos em modelos de decisão**. In: CONTINI, E.; ARAÚJO, J.D.; OLIVEIRA, A.J. et al. Planejamento da propriedade agrícola, modelos de decisão. Brasília, EMBRAPA - DDT, 1984. p.237-259.

DASKALAKI,C; SKIADOPOULOS,G., Should investors include *commodities* in their portfolios after all? New evidence. **Journal of Banking & Finance**, V.35 , P. 2606–2626, 2011.

DIAS, C.T.S., **Planejamento de uma fazenda em condições de risco: Programação Linear e Simulação Multidimensional**. 113p. Tese de Doutorado em Estatística e Experimentação Agronômica. ESALQ-USP, São Paulo, 1996.

DIAS, C.T.S.; SAMARANAYAKA, A.; MANLY, B. On the use of correlated beta random variables with animal population modeling, **Ecological modeling**, v.215, p.293-300, 2008.

DILLON, J.L. **Avaliação de tecnologias agrícolas alternativas sob risco**. Fortaleza, Centro de Ciências Agrárias, Universidade Federal do Ceará, 1975, 22 p.

DRIESSEN, J.; LAEVEN, L. International portfolio diversification benefits: Cross-country evidence from a local perspective. **Journal of Banking & Finance** v.31, p. 1693–1712, 2007.

EMBRAPA, **Boletim de Pesquisa Online Nº 2, dez./99**, Disponível em: <http://www.cnpt.embrapa.br/biblio/p_bo02_3.htm> . Acesso em 13 março de 2011.

FAN, X. et al. **SAS[®] for Monte Carlo Studies: a guide for quantitative researchers**. Cary, NC, USA: SAS institute inc., 2001.

HAZELL, P.B. A linear alternative to quadratic and semvariance programming for farm planning under uncertainty. **Amer. J. Agri. Econ.** V.53, n.1, p. 53-62, 1971.

KNIGHT, F.H. **Risk, uncertainty and profit**. Boston, Houghton – Mifflin, 1921.

LAW, A., M. **Simulation and modeling and analysis**. 4^a ed. Nova York: McGraw Hill, 2007, 768p.

MARKOWITZ, H. Portfolio selection. **Journal of Finance**, v.7, n.1, p. 71-91, 1952.

MARKOWITZ, H. **Portfolio selection: Efficiency Diversification Of Investments**. 1^a ed. New York: John Wiley, 1959.

MARQUES, P.A.A. FRIZZONE, J.A.; CAIXETA-FILHO, J.V., Estudo de ocupação econômica em área agrícola na região de Piracicaba-SP, incluindo risco através de programação linear. **Biosci.J.**, Uberlândia, v.25, n.1, p.30-41, Jan/Fev, 2009.

MATTUELLA L.J. Seleção de planos de produção para pequenos produtores agrícolas. **Análise Econômica**. Porto Alegre, UFRGS, nº 14, nov, 1990.

METROPOLIS, N.; ULAN, S. The Monte carlo Method. **Journal of the American Statistical Association**, V.44, p. 335-341, 1949.

MOSS CHALES B. **Risk, uncertainty and the agricultural firm**. USA: World Scientific Publishing Co. Pte. Lt., 2010, 292p.

MOUTINHO, D.A.; SANDERS JUNIOR, J.H.; WEBER, M.T. Tomada de decisão sob condições de risco em relação à nova tecnologia para a produção de feijão de corda. **Revista de Economia Rural**. Brasília, v.16, n.4, p.41-58, 1978.

PARIKH, A.; SHAH K.M. Measurement of cost inefficiency with safety first measure of risk. **Agricultural Economics** v.1, n.11, p. 197-206, 1994.

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

PECHE FILHO, A. **A tomada de decisão na escolha de sistemas de preparo do solo**. 2007. Disponível em: <http://www.infobibos.com/Artigos/2007_1/decisao/Index.htm> , Acesso em 01 abril 2010.

PÉRES, A.R.P. **Baixa produtividade do milho como consequência da tomada de decisão sob condições de risco**. Piracicaba, SP, 1981. 118p. Dissertação de Mestrado ESALQ/USP.

PÉRES, F.C. Planejamento da empresa agrícola em condições de risco. In: CONTINI, E.; ARAÚJO, J.D.; OLIVEIRA, A.J. et al. **Planejamento da propriedade agrícola, modelos de decisão**. Brasília, EMBRAPA - DDT, 1984, p. 273-287.

PRATT, JOHN W. Risk aversion in the small and in the large. **Econometrica**, v.32, n.1-2, 122-136, 1964.

RAYCHAUDHURI, S., **Introduction to Monte Carlo simulation**. In: WINTER SIMULATION CONFERENCE, 2008.

ROMAN, D.; DARBY, K.; MITRA, G., Portfolio construction based on stochastic dominance and target return distributions. **Math. Program.**, Ser. B 108, p.541-569, 2006.

SALIMONU, K.K.; FALUSI, A.O.; OKURUWA, V.O.; YUSURF, S.A., Modeling Efficient Resource Allocation Patterns for Food Crop Farmers in Nigeria: An Application of T- MOTAD Analysis. **International Journal of Agricultural Economics & Rural Development** n.1, v.1, p. 2008.

SALTELLI, A., **Global Sensitivity Analysis: An Introduction**, Sensitivity analysis of model output, Kenneth M. Hanson and François M. Hemez, eds., Los Alamos National Laboratory, 2005

SALTELLI, A. Sensitivity Analysis for Importance Assessment. **Risk Analysis**, vol.22, n.3, p.579-590, 2002.

SANT'ANNA A.C., **O uso econômico da reserva legal no cerrado: Uma simulação do extrativismo sustentável do pequi**. Dissertação de Mestrado, ESALQ/USP, Piracicaba, SP, 2011, 131p.

SILVA, L.M.R., **Rentabilidade em condições de incerteza de uma média empresa produtora de frango de corte no estado do Ceará**, Revista econômica do nordeste, Fortaleza, v.30, n.4, p. 1000-1013, out-dez, 1999.

SOARES, F.B. Programação da empresa em condições de incerteza: Uma possível abordagem. **Economia**, v.1, n.2, p.319-35, maio, 1977.

TAUER, L., Target-MOTAD. **American Journal Agriculture Economics**, v.65, n.3, p.606-610, 1983.

TURVEY C.G.; ESCALANTE C.L.; NGANJE, W., Developments in portfolio management and risk programming techniques for agriculture. **Agriculture Finance Review**, p. 219-244, 2005.

WAGNER S. A. L., **Novas e velhas abordagens de teoria da escolha e da utilidade**. Dissertação de Mestrado em Economia. PUCRS, Rio Grande do Sul, 2009. 179 p.

WALLS, M.R. Combining decision analysis and portfolio management to improve project selection in the exploration and production firm. **Journal of Petroleum Science and Engineering**, v.44, n.1-2, p. 55-65, outubro, 2004.

YANG, J. Convergence and uncertainty in Monte Carlo based sensitivity analysis. **Environmental Modeling & Software**, vol.26, n.4, p. 444-457, 2011.

APÊNDICE A: COEFICIENTE DE CORRELAÇÃO ENTRE AS COMMODITIES ESCOLHIDAS

	t1	t2	t3	t4	t5	t6	t7
t1	1	-0.22	0.33	0.65	-0.24	0.55	0.33
t2	-0.22	1	0.5	-0.22	0.77	-0.25	0.17
t3	0.33	0.5	1	0.31	0.28	0.44	0.44
t4	0.65	-0.22	0.31	1	-0.35	0.66	0.34
t5	-0.24	0.77	0.28	-0.35	1	-0.27	0.07
t6	0.55	-0.25	0.44	0.66	-0.27	1	0.61
t7	0.33	0.17	0.44	0.34	0.07	0.61	1

APÊNDICE B: PROGRAMA PARA SIMULAR COEFICIENTES DEPENDENTES E SOLUCIONAR O PROBLEMA DE PROGRAMAÇÃO LINEAR

```
title "Geração de 7 variáveis com distribuição triangular seguindo o
modelo de MARKOWITZ";
```

```
proc iml ;
```

```
wrksize=100000;
```

```
K=7;
```

```
N=1000;
```

```
M={0 0 0 0 0 0 0};
```

```
S={1      -0.22      0.33      0.65  -0.24  0.55  0.33,
    -0.22      1      0.5      -0.22  0.77  -0.25  0.17,
    0.33      0.5      1      0.31  0.28  0.44  0.44,
    0.65      -0.22      0.31      1      -0.35  0.66
    0.34,
    -0.24      0.77      0.28      -0.35  1      -0.27  0.07,
    0.55      -0.25      0.44      0.66  -0.27  1      0.61,
    0.33      0.17      0.44      0.34  0.07  0.61  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;

Z=X`;

varnames='X1':'X7';

create NOVO from Z[colname=varnames];
append from Z;

quit;

data PERIME ; set NOVO;
options ps=66 ls=75;

* Definição dos parâmetros da distribuição triangular;

a1=64.1;   m1=95.1;   b1=359.9;
a2=9.2;    m2=82.3;   b2=127.5;
a3=40.5;   m3=46.8;   b3=399.7;
a4=68.9;   m4=157.6;  b4=552.7;
a5=17.2;   m5=103.4;  b5=153.8;
a6=65.3;   m6=77.4;   b6=150.7;
a7=66;     m7=95.8;   b7=385.5;

* Geração de coeficientes com distribuição uniforme;

y1=erf(sqrt(2)*x1/2)/2+0.5;
y2=erf(sqrt(2)*x2/2)/2+0.5;
y3=erf(sqrt(2)*x3/2)/2+0.5;
y4=erf(sqrt(2)*x4/2)/2+0.5;
y5=erf(sqrt(2)*x5/2)/2+0.5;
y6=erf(sqrt(2)*x6/2)/2+0.5;
y7=erf(sqrt(2)*x7/2)/2+0.5;

* Geração de coeficientes 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-m3)*(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));

if (y7>=0) and (y7<=(m7-a7)/(b7-a7))
  then t7=sqrt((b7-a7)*(m7-a7)*y7)+a7;
if (y7>(m7-a7)/(b7-a7)) and (y7<=1)
  then t7= b7-sqrt((y7-1)*(b7-m7)*(a7-b7));

*****Criação de variáveis a partir do Modelo de Markowitz*****;

d1=((t1-165.40)/4082.37)+((t2-65.69)/-300.73)+((t3-
121.39)/1071.08)+
((t4-234.94)/2978.21)+((t5-90.86)/-383.54)+((t6-97.99)/770.57)+((t7-
180.81)/1182.49);

d2=((t1-165.40)/-300.73)+((t2-65.69)/1269.44)+((t3-
121.39)/997.19)+
((t4-234.94)/1284.48)+((t5-90.86)/815.38)+((t6-97.99)/82.42)+((t7-
180.81)/284.91);

d3=((t1-165.40)/1071.08)+((t2-65.69)/997.19)+((t3-
121.39)/3044.05)+
((t4-234.94)/3569.35)+((t5-90.86)/465.43)+((t6-97.99)/558.57)+((t7-
180.81)/1144.19);

d4=((t1-165.40)/2978.21)+((t2-65.69)/1284.48)+((t3-
121.39)/3569.35)+
((t4-234.94)/9476.82)+((t5-90.86)/-355.52)+((t6-97.99)/1075.23)+((t7-
180.81)/2049.93);

d5=((t1-165.40)/-383.54)+((t2-65.69)/815.38)+((t3-
121.39)/465.43)+

```

```

((t4-234.94)/-355.52)+((t5-90.86)/1240.8)+((t6-97.99)/-71.77)+((t7-
180.81)/98.84);

d6=((t1-165.40)/770.57)+((t2-65.69)/82.42)+((t3-
121.39)/558.57)+
((t4-234.94)/2049.93)+((t5-90.86)/-71.77)+((t6-97.99)/431.72)+((t7-
180.81)/768.53);

d7=((t1-165.40)/1182.49)+((t2-65.69)/284.91)+((t3-
121.39)/1144.19)+
((t4-234.94)/2049.93)+((t5-90.86)/98.84)+((t6-97.99)/768.53)+((t7-
180.81)/2171.69);

title 'Impressão das variáveis geradas com distribuição Triangular';
proc print data=PERIME;
options nodate nostimer nonumber;
VAR t1 t2 t3 t4 t5 t6 t7;
run;

title 'Coeficiente de correlação das variáveis com geradas com
distribuição triangular';
proc corr data=PERIME;
VAR t1 t2 t3 t4 t5 t6 t7;
with t1 t2 t3 t4 t5 t6 t7;
run;

title 'Variáveis Criadas a partir do modelo de Markowitz';
proc print data=PERIME;
var d1 d2 d3 d4 d5 d6 d7;
run;

title 'SOLUCAO DO PROBLEMA DE PROGRAMACAO LINEAR';

data IRRIGA;
input _id_ $ y1 y2 y3 y4 y5 y6 y7 _type_ $ _rhs_ @@;
cards;
OBJET      0.1744    0.4274    0.4131    0.1633    -0.603    0.9783
0.6504      MIN      .
REST_ORÇA  1          1          1          1          1          1
1          EQ      100
;
proc lp data=IRRIGA;
run;
%macro AUX ;
data RESULTS; stop; run;
%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_
y1 y2 y3 y4 y5 y6 y7
_type_ _rhs_;
retain d1_ d2_ d3_ d4_ d5_ d6_ d7_;
if _n_ eq 1 then do;
d1_=d1;d2_=d2;d3_=d3;d4_=d4;d5_=d5;d6_=d6;d7_=d7;
end;
d1=d1_ d2=d2_ d3=d3_ d4=d4_ d5=d5_ d6=d6_ d7=d7_;
```



```

if _id_ eq 'OBJET' then do;
  y1=d1; y2=d2; y3=d3; y4=d4; y5=d5; y6=d6; y7=d7;
end;

run;
proc print;run;
proc LP data=LPAUX primalout=PP;

data TESTA;
set PP; keep Z p1 p2 p3 p4 p5 p6 p7;
if _var_='OBJET' then Z=_value_;
if _var_='y1' then p1=_value_;
if _var_='y2' then p2=_value_;
if _var_='y3' then p3=_value_;
if _var_='y4' then p4=_value_;
if _var_='y5' then p5=_value_;
if _var_='y6' then p6=_value_;
if _var_='y7' then p7=_value_;

if Z='.' and p1='.' and p2='.' and p3='.' and p4='.' and p5='.'
and p6='.' and p7='.'
then delete;output;
proc means data=TESTA nway noprint missing;
var Z p1 p2 p3 p4 p5 p6 p7;
output out=TESTAR
min(Z p1 p2 p3 p4 p5 p6 p7)=Z p1 p2 p3 p4 p5 p6 p7;
run;
proc append base=EMPTY data=TESTAR;
run;
%end;
%mend aux;
%aux;
TITLE 'IMPRESSÃO DOS RESULTADOS OBTIDOS PELA PROGRAMAÇÃO LINEAR';
proc print data=EMPTY;
var Z;
run;
TITLE 'DISTRIBUIÇÃO DE FREQUÊNCIAS DOS RESULTADOS OBTIDOS PELA
PROGRAMAÇÃO LINEAR';
proc freq data=EMPTY;
run;
proc univariate data=EMPTY normal;

var Z; run;

```

APÊNDICE C: PROGRAMA PARA SIMULAR COEFICIENTES INDEPENDENTES E SOLUCIONAR O PROBLEMA DE PROGRAMAÇÃO LINEAR

```

title "Geração de 7 variáveis com distribuição triangular seguindo o
modelo de MARKOWITZ";

proc iml ;

wrksize=100000;
K=7;

```

```

N=1000;
M={0 0 0 0 0 0 0};
S={1      0      0      0      0      0      0      0,
    0      1      0      0      0      0      0      0,
    0      0      1      0      0      0      0      0,
    0      0      0      1      0      0      0      0,
    0      0      0      0      1      0      0      0,
    0      0      0      0      0      1      0      0,
    0      0      0      0      0      0      1      0,
    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;
Z=X`;

varnames='X1':'X7';

create NOVO from Z[colname=varnames];
append from Z;

quit;

data PERIME ; set NOVO;
options ps=66 ls=75;

* Definição dos parâmetros da distribuição triangular;

a1=64.1;   m1=95.1;   b1=359.9;
a2=9.2;    m2=82.3;   b2=127.5;
a3=40.5;   m3=46.8;   b3=399.7;
a4=68.9;   m4=157.6;  b4=552.7;
a5=17.2;   m5=103.4;  b5=153.8;
a6=65.3;   m6=77.4;   b6=150.7;
a7=66;     m7=95.8;   b7=385.5;

* Geração de coeficientes com distribuição uniforme;

y1=erf(sqrt(2)*x1/2)/2+0.5;
y2=erf(sqrt(2)*x2/2)/2+0.5;
y3=erf(sqrt(2)*x3/2)/2+0.5;
y4=erf(sqrt(2)*x4/2)/2+0.5;
y5=erf(sqrt(2)*x5/2)/2+0.5;
y6=erf(sqrt(2)*x6/2)/2+0.5;
y7=erf(sqrt(2)*x7/2)/2+0.5;

```

```

* Geração de coeficientes 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-m3)*(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));

if (y7>=0) and (y7<=(m7-a7)/(b7-a7))
  then t7=sqrt((b7-a7)*(m7-a7)*y7)+a7;
if (y7>(m7-a7)/(b7-a7)) and (y7<=1)
  then t7= b7-sqrt((y7-1)*(b7-m7)*(a7-b7));

*****Criação de variáveis a partir do Modelo de Markowitz*****;

d1=((t1-165.40)/4082.37)+((t2-65.69)/-300.73)+((t3-
121.39)/1071.08)+
((t4-234.94)/2978.21)+((t5-90.86)/-383.54)+((t6-97.99)/770.57)+((t7-
180.81)/1182.49);

d2=((t1-165.40)/-300.73)+((t2-65.69)/1269.44)+((t3-
121.39)/997.19)+
((t4-234.94)/1284.48)+((t5-90.86)/815.38)+((t6-97.99)/82.42)+((t7-
180.81)/284.91);

d3=((t1-165.40)/1071.08)+((t2-65.69)/997.19)+((t3-
121.39)/3044.05)+
((t4-234.94)/3569.35)+((t5-90.86)/465.43)+((t6-97.99)/558.57)+((t7-
180.81)/1144.19);

d4=((t1-165.40)/2978.21)+((t2-65.69)/1284.48)+((t3-
121.39)/3569.35)+

```

```

((t4-234.94)/9476.82)+((t5-90.86)/-355.52)+((t6-97.99)/1075.23)+((t7-
180.81)/2049.93);

d5=((t1-165.40)/-383.54)+((t2-65.69)/815.38)+((t3-
121.39)/465.43)+
((t4-234.94)/-355.52)+((t5-90.86)/1240.8)+((t6-97.99)/-71.77)+((t7-
180.81)/98.84);

d6=((t1-165.40)/770.57)+((t2-65.69)/82.42)+((t3-
121.39)/558.57)+
((t4-234.94)/2049.93)+((t5-90.86)/-71.77)+((t6-97.99)/431.72)+((t7-
180.81)/768.53);

d7=((t1-165.40)/1182.49)+((t2-65.69)/284.91)+((t3-
121.39)/1144.19)+
((t4-234.94)/2049.93)+((t5-90.86)/98.84)+((t6-97.99)/768.53)+((t7-
180.81)/2171.69);

title 'Impressão das variáveis geradas com distribuição Triangular';
proc print data=PERIME;
options nodate nostimer nonumber;
VAR t1 t2 t3 t4 t5 t6 t7;
run;

title 'Coeficiente de correlação das variáveis com geradas com
distribuição triangular';
proc corr data=PERIME;
VAR t1 t2 t3 t4 t5 t6 t7;
with t1 t2 t3 t4 t5 t6 t7;
run;

title 'Variáveis Criadas a partir do modelo de Markowitz';
proc print data=PERIME;
var d1 d2 d3 d4 d5 d6 d7;
run;
title 'ANÁLISE DE RESULTADOS APÓS AS SIMULAÇÕES';

title 'SOLUCAO DO PROBLEMA DE PROGRAMACAO LINEAR';

data IRRIGA;

input _id_ $ y1 y2 y3 y4 y5 y6 y7 _type_ $ _rhs_ @@;
cards;
OBJET      0.1744    0.4274    0.4131    0.1633    -0.603    0.9783
0.6504      MIN      .
REST_ORÇA  1          1          1          1          1          1
1          EQ      100
;

proc lp data=IRRIGA;
run;

%macro AUX ;
data RESULTS; stop; run;
%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_
    y1 y2 y3 y4 y5 y6 y7
    _type_ _rhs_;
retain d1_ d2_ d3_ d4_ d5_ d6_ d7_;
if _n_ eq 1 then do;
d1_=d1;d2_=d2;d3_=d3;d4_=d4;d5_=d5;d6_=d6;d7_=d7;
end;
d1=d1_; d2=d2_; d3=d3_;d4=d4_;d5=d5_;d6=d6_;d7=d7_;
if _id_ eq 'OBJET' then do;
y1=d1; y2=d2; y3=d3; y4=d4; y5=d5; y6=d6; y7=d7;
end;

run;
proc print;run;
proc LP data=LPAUX primalout=PP;

data TESTA;
set PP; keep Z p1 p2 p3 p4 p5 p6 p7;
if _var_='OBJET' then Z=_value_;
if _var_='y1' then p1=_value_;
if _var_='y2' then p2=_value_;
if _var_='y3' then p3=_value_;
if _var_='y4' then p4=_value_;
if _var_='y5' then p5=_value_;
if _var_='y6' then p6=_value_;
if _var_='y7' then p7=_value_;

if Z='.' and p1='.' and p2='.' and p3='.' and p4='.' and p5='.'
and p6='.' and p7='.'
then delete;output;

proc means data=TESTA nway noprint missing;
var Z p1 p2 p3 p4 p5 p6 p7;
output out=TESTAR
min(Z p1 p2 p3 p4 p5 p6 p7)=Z p1 p2 p3 p4 p5 p6 p7;
run;
proc append base=EMPTY data=TESTAR;

run;
%end;
%mend aux;

%aux;

TITLE 'IMPRESSÃO DOS RESULTADOS OBTIDOS PELA PROGRAMAÇÃO LINEAR';
proc print data=LPAUX;
run;

TITLE 'DISTRIBUIÇÃO DE FREQUÊNCIAS DOS RESULTADOS OBTIDOS PELA
PROGRAMAÇÃO LINEAR';
proc freq data=EMPTY;
run;

```

```
proc univariate data=EMPTY normal;  
var Z; run;
```