

NICHOLAS CARBONE

**ANÁLISE DE INVESTIMENTO NO MERCADO DE REAL ESTATE
NORTE AMERICANO UTILIZANDO UMA ABORDAGEM
ESTOCÁSTICA PELO MÉTODO DE MONTE CARLO.**

Monografia apresentada à Escola
Politécnica da Universidade de São Paulo
como requisito para a obtenção do título de
MBA em Engenharia Financeira.

São Paulo

2018

NICHOLAS CARBONE

**ANÁLISE DE INVESTIMENTO NO MERCADO DE REAL ESTATE
NORTE AMERICANO UTILIZANDO UMA ABORDAGEM
ESTOCÁSTICA PELO MÉTODO DE MONTE CARLO.**

Monografia apresentada à Escola
Politécnica da Universidade de São Paulo
como requisito para a obtenção do título de
MBA em Engenharia Financeira.

Área de Concentração: Engenharia
Financeira

Orientador: Prof. Dr. André Cury Maialy.

São Paulo

2018

Dedico este trabalho a minha
esposa Silviane e minhas filhas
Luiza e Sofia.

AGRADECIMENTO

Agradeço a Deus por ter me dado todo suporte necessário para trabalhar, estudar e correr atrás dos meus sonhos. Agradeço a minha esposa Silviane e minhas filhas Luiza e Sofia pelo apoio e carinho que sempre recebi. Agradeço ao professor André Maialy pela competente orientação e paciência. Agradeço a todos os professores do curso de Engenharia Financeira pelos valiosos ensinamentos.

Tal como a chuva caída
Fecunda a terra, no estio,
Para fecundar a vida
O trabalho se inventou.
Feliz quem pode, orgulhoso, Dizer:
"Nunca fui vadio:
E, se hoje sou venturoso,
Devo ao trabalho o que sou!"
É preciso, desde a infância,
Ir preparando o futuro;
Para chegar à abundância
É preciso trabalhar.
Não nasce a planta perfeita,
Não nasce o fruto maduro;
E, para ter a colheita,
É preciso semear...

O trabalho: Olavo Bilac

RESUMO

Atualmente a utilização de métodos tradicionais como FCD, múltiplos, bem como a consideração de cenários desfavoráveis, otimistas ou intermediários tem sido a forma utilizada por alguns profissionais do mercado para a análise e modelagem de investimentos e riscos, especialmente investimentos ou projetos fora da fronteira do mercado financeiro. Entretanto, estas metodologias são incapazes de modelar as diversas incertezas existentes nestas modalidades, impossibilitando assim uma correta tomada de decisão a respeito destes investimentos.

Neste trabalho será apresentado uma aplicação de modelagem estocástica utilizando o método de Monte Carlo para análise financeira da atividade de *flip* no mercado de *real estate* norte americano, mais especificamente, na região leste de New Jersey. O modelo considera as diversas variáveis aleatórias que compõem a atividade e apresenta um completo mapeamento das incertezas, riscos e retornos esperados. Além disso, o trabalho apresenta com base nas saídas do modelo, formas de maximizar retornos minimizando riscos. Possibilitando uma melhor forma de gerenciar e controlar o investimento pela maior compreensão dos riscos e benefícios envolvidos na atividade.

Palavras-chave: Engenharia Financeira. Método de Monte Carlo. Mercado de Real Estate.

ABSTRACT

Nowadays use of traditional methods such as DCFs, multiples as well as the consideration of unfavorable, optimistic and intermediate scenarios has been the way used by some market professionals to analyse and model investments and risk, especially investments and projects outside the financial market border. However, these methodologies are unable to model various uncertainties that are in these modalities, making it impossible to correctly make decisions about these investments.

This work will present a stochastic modeling application using the Monte Carlo simulation for financial analysis of the flip activity in the North American Real Estate market, more specifically in the east of New Jersey region. The model considers various random variables that compose the activity and presents a complete mapping of the uncertainties, risks and expected returns. In addition, the work presents, based on the outputs of the model, ways to maximize returns while minimizing risks. Enabling a better way to manage and control the investment by better understanding of the risks and benefits involved in the activity.

Keywords: Financial Engineering. Monte Carlo Simulation. Real Estate Market.

SUMÁRIO

1. Introdução	1
1.1. Motivação do trabalho	1
1.2 Objetivo e justificativa do trabalho	2
2. Revisão Bibliográfica	3
2.1 Simulação numérica de Monte Carlo	3
2.1.1 Formulação Básica do Método de Monte Carlo	3
2.1.2 Exemplos Básicos do Método de Monte Carlo	4
2.2 Distribuições de probabilidades	5
2.2.1 Distribuição Binomial	6
2.2.2 Distribuição Normal	6
2.2.3 Distribuição Exponencial	7
2.2.4 Distribuição Gama	8
2.2.5 Distribuição Beta	8
2.3 Gerenciamento de risco	9
2.3.1 Riscos e Incertezas	9
2.3.2 Natureza do Risco	11
2.3.3 A Estatística do Risco	12
2.3.3.1 Medindo o Centro da Distribuição– 1º Momento	13
2.3.3.2 Medindo o Spread da Distribuição– 2º Momento	13
2.3.3.3 Medindo a Inclinação da Distribuição– 3º Momento	15
2.3.3.4 Medindo os Efeitos de calda Catastróficos– 4º Momento	16
2.3.4 Medindo Risco	17
2.3.5 Simulação de Monte Carlo para Análise de Risco	19
2.4 Utilização do método de Monte Carlo no mercado imobiliário	20
3. Desenvolvimento do Modelo	22
3.1 Aprofundamento da caracterização do objeto de estudo	22
3.2 Abordagem do objeto de estudo	23
3.3 Modelagem	24
3.4 Obtenção dos resultados	27
4. Aplicação da Teoria em um Caso Específico	29
4.1 Coleta e tratamento estatístico da amostra	29

4.2 Fit das distribuições de probabilidade do modelo	32
4.2.1 Fit da distribuição da variável X_1	32
4.2.2 Fit da distribuição da variável X_2	35
4.2.3 Fit da distribuição da variável X_3	38
4.2.4 Fit da distribuição da variável X_4	41
4.3 Simulação de Monte Carlo	44
5. Apresentação e Discussão dos Resultados	49
6. Conclusões	51
7. Trabalhos Futuros	53
Referências Bibliográficas	54

1. Introdução

1.1 Motivação do trabalho

Atualmente, parte das análises financeiras de projetos e investimentos são realizadas por metodologias incapazes de modelar e analisar os riscos existentes nestas atividades de investimento. Metodologias como análises por múltiplos, fluxo de caixa descontado (FCD), entre outros métodos são comumente usados pelas entidades de ensino e instituições financeiras para avaliar projetos e investimentos financeiros.

Infelizmente, tais metodologias utilizam muitas suposições e arbitrariedades em suas análises, desprezando os riscos e incertezas existentes. Nestes métodos, essas incertezas geralmente são modeladas com a construção de apenas alguns cenários: cenário otimista, cenário pessimista e cenário intermediário, os quais claramente não exemplificam o mundo real onde os eventos futuros são altamente imprevisíveis.

Assim sendo, a utilização de um modelo convencional de FCD, análises de múltiplos e análises de cenários, podem ser um bom início, entretanto, existem formas mais adequadas e matematicamente mais sofisticadas para analisar e gerenciar riscos e incertezas em investimentos e projetos. Em termos quantitativos, modelos estocásticos e determinísticos são totalmente opostos pois os modelos estocásticos são caracterizados por uma evolução aleatória. Enquanto os modelos determinísticos são baseados na suposição de uma simples forma de como o processo evolui durante o tempo. Modelos estocásticos assumem que um imprevisível componente e alguns tipos de aleatoriedade irão dirigir futuramente um conjunto de eventos. Este elemento aleatório implica que mesmo conhecendo a situação inicial existirá várias possibilidades de evolução do processo no futuro.

O método de Monte Carlo foi desenvolvido na década de 30 pelos físicos Enrico Fermi, John Von Neumann e Stanislaw Ulan. Inicialmente, o método foi aplicado para o estudo da difusão aleatória de nêutrons da bomba atômica. A utilização do método de Monte Carlo no cenário financeiro, para a análise e avaliação de projetos e investimentos, tem sido alvo de várias pesquisas pela sua capacidade de poder modelar as várias aleatoriedades envolvidas em um projeto.

Segundo Moore e Weatherford (2006), a simulação de Monte Carlo pode ser amplamente utilizada na avaliação de projetos auxiliando na tomada de decisão.

1.2 Objetivo e justificativa do trabalho

Este trabalho faz parte das disciplinas obrigatórias para obtenção do título de MBA exigido pelo Programa de Educação Continuada da Escola Politécnica da USP. O objetivo é apresentar um modelo estocástico utilizando o método de Monte Carlo para a análise financeira da atividade de flip no mercado de real estate norte americano. O foco do trabalho será a identificação das incertezas existentes na atividade de flip e na construção de um modelo matemático que possibilite maior controle e gerenciamento sobre os riscos existentes nesta atividade de investimento.

A atividade de flip no mercado de real estate norte americano consiste basicamente na aquisição de um imóvel, realização de reforma ou renovação e por fim, a comercialização deste imóvel.

Segundo Pustejovsky (2017), a atividade de flip com imóveis residenciais teve seu auge após a crise imobiliária de 2007-2008 onde a oferta por imóveis com baixíssimos preços era muito grande. Geralmente esses imóveis apresentavam avarias geradas pela má conservação e abandono, o que se tornou uma grande oportunidade de investimento neste segmento. Com o início do reaquecimento da economia em 2011, a aquisição deste tipo de imóveis por meio de leilões ou mesmo *foreclosure* (compras diretas com o banco) virou prática disputada entre os investidores do setor e uma solução para as instituições financeiras liquidarem seus títulos.

Ainda segundo Pustejovsky (2017), com o significativo aumento da competitividade nesta atividade de investimento e com a redução da oferta dos imóveis, devido ao reaquecimento da economia norte americana, ocorreu a redução significativa dos retornos, demandando maior cautela do investidor.

Assim sendo, a utilização de uma metodologia matematicamente mais sofisticada auxiliaria o investidor na correta tomada de decisão em relação a este investimento, minimizando riscos e maximizando retornos.

2. Revisão Bibliográfica

2.1 Simulação Numérica de Monte Carlo.

Segundo Wang (2012), o método de Monte Carlo é uma ferramenta muito flexível e poderosa para estimar integrais e valores esperados. Grande parte das análises quantitativas em finanças ou análises de gerenciamento de riscos frequentemente envolvem cálculos para determinação de valores esperados de uma função de variáveis aleatórias, sendo o método de Monte Carlo amplamente utilizado para solução destes problemas na indústria financeira.

2.1.1 Formulação Básica do Método de Monte Carlo

Considere uma situação hipotética de estimação do valor esperado de uma função de uma variável aleatória X :

$$\mu = E[h(X)]. \quad (1)$$

O procedimento do método de Monte Carlo pode ser basicamente dividido em duas etapas:

1. Geração de amostras ou variáveis aleatórias independentes e identicamente distribuídas (iid) X_1, X_2, \dots, X_n , que tenham a mesma distribuição que X .
2. A estimativa do valor esperado μ é definido para ser a média da amostra.

$$\hat{\mu} = \frac{1}{n} [h(X_1) + h(X_2) + \dots + h(X_n)]. \quad (2)$$

Como em todo método numérico, faz-se necessário uma discretização adequada para minimização dos erros. Neste caso, uma maior discretização significa um maior número de indivíduos (n).

2.1.2 Exemplos Básicos do Método de Monte Carlo

A seguir, é apresentado um exemplo da utilização do método de Monte Carlo para estimar o preço de uma opção de compra (CALL) pela suposição que o preço da ação é um movimento Browniano geométrico.

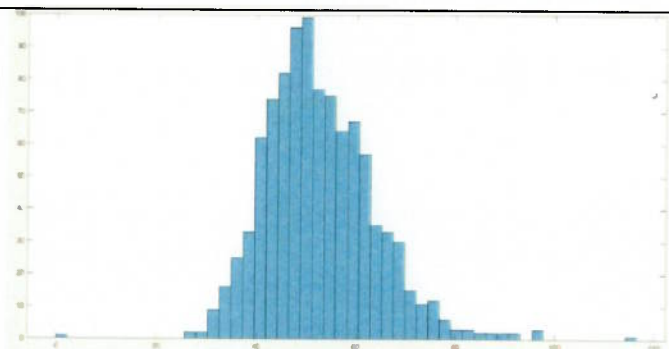
Considera-se que o preço da opção de compra com strike K e vencimento T é dada por: $v = E[e^{-rT}(S_T - K)^+]$, onde r é a taxa livre de risco. O valor esperado é obtido sob a medida de probabilidade neutra ao risco, onde o preço da ação é um movimento Browniano: $S_T = S_0 \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)T + \sigma W_T\right\}$.

Posteriormente, realiza-se a geração de uma amostra com n elementos (simulações) do *payoff* da opção. Sabendo que W_T é normalmente distribuído com média 0 e variância T (característica do movimento Browniano), pode-se escrever $W_T = z\sqrt{T}$ para uma variável aleatória normal padrão z .

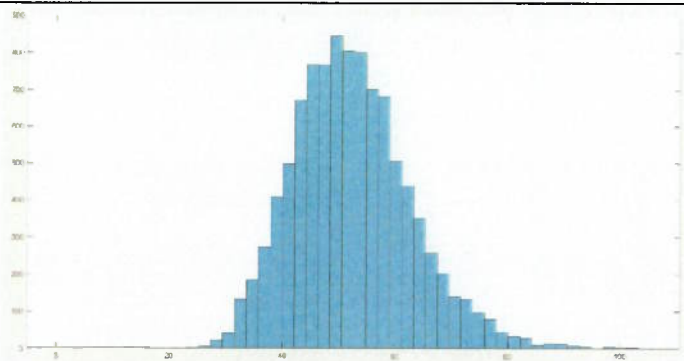
Código MatLab

```
for i = 1:n
    St(i+1) = So*exp(((r-(0.5*(sigma^2))))*T + normrnd(0,1)*sigma*(T)^0.5);
    if St(i)<K
        V(i)=0;
    else
        V(i)=exp(-r*T)*(St(i)-K);
    end
end
Av=mean(V)
```

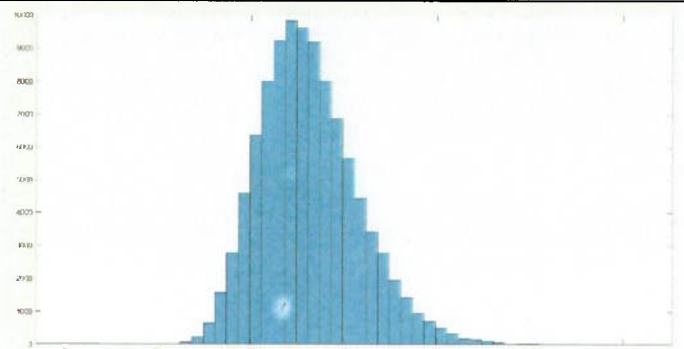
Resultados com n=1.000 simulações

Características	Distribuição do Preço da ação S_T
$K=40$ $S_0=50$ $\sigma=0,2$ $T=1$ (Tf-t) $r=5\%$ $v = 12,6048$ $v_a=12,2944$ (v _a : valor teórico, BS)	

Resultados com $n=10.000$ simulações

Características	Distribuição do Preço da ação S_T
$K=40$ $S_0=50$ $\sigma=0,2$ $T=1$ (T_f-t) $r=5\%$ $v = 12,2292$ $v_a=12,2944$ $(v_a: \text{valor teórico, BS})$	

Resultados com $n=100.000$ simulações

Características	Distribuição do Preço da ação S_T
$K=40$ $S_0=50$ $\sigma=0,2$ $T=1$ (T_f-t) $r=5\%$ $v = 12,2860$ $v_a=12,2944$ $(v_a: \text{valor teórico, BS})$	

Observou-se uma queda do erro na ordem de $1/\sqrt{n}$ do valor da call estimado pelo método de Monte Carlo (v) em comparação com o valor teórico (v_a) encontrado pela formulação de Black-Scholes.

2.2 Distribuição de Probabilidades

Segundo Navidi (2012), a inferência estatística envolve a extração de uma amostra de uma população e a análise dos dados da amostra para entender a população. Em muitas situações, tem-se algum conhecimento sobre a função de massa de probabilidade ou sobre a função densidade de probabilidade da população. Nestes casos, essas funções muitas vezes podem ser bem aproximadas por uma das várias famílias padrão de curvas ou funções. A seguir, são apresentadas algumas destas famílias padrão, descrevendo sob quais condições elas são mais apropriadas.

2.2.1 Distribuição Binomial

Considere um experimento que pode resultar em um entre dois resultados. Um resultado é denominado “sucesso” e o outro de “insucesso”. A probabilidade de sucesso é indicada por p . Portanto a probabilidade de insucesso é $1-p$. Esses ensaios são denominados ensaios de Bernoulli com probabilidade e sucesso p .

Muitos experimentos consistem em realizar uma sequência de n ensaios de Bernoulli. Então, considerando estes ensaios como sendo ensaios independentes, o número de sucessos é uma variável aleatória, a qual se diz ter uma distribuição Binomial. Logo, diz-se que X tem uma distribuição Binomial com parâmetros n e p onde X é o número de sucessos nos n ensaios. Utiliza-se a notação $X \sim \text{Bin}(n,p)$ (NAVIDI, 2012).

Se $X \sim \text{Bin}(n,p)$, a função de massa de probabilidade de X é:

$$p(x) = P(X = x) = \begin{cases} \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} & x = 0, 1, \dots, n \\ 0 & \text{caso contrário} \end{cases} \quad (3)$$

Se $X \sim \text{Bin}(n,p)$, então a média e a variância de X são dadas por:

$$\mu_X = np \quad (4)$$

$$\sigma_X^2 = np(1-p) \quad (5)$$

2.2.2 Distribuição Normal

Segundo Navidi (2016), distribuição normal ou distribuição gaussiana é de longe a distribuição mais usada em estatística. Essa distribuição fornece um bom modelo para muitas, embora não todas, populações que possuem variáveis aleatórias contínuas. A média de uma variável aleatória normal pode ter qualquer valor e a variância pode ter qualquer valor positivo. A função densidade de probabilidade de uma variável aleatória normal com média μ e variância σ^2 é dada por:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)} \quad (6)$$

Se X for uma variável aleatória cuja função densidade de probabilidade seja normal com média μ e variância σ^2 , escrevemos $X \sim N(\mu, \sigma^2)$.

2.2.3 Distribuição Exponencial

A distribuição exponencial é uma distribuição contínua que algumas vezes é usada para modelar o tempo que decorre antes do evento acontecer. Geralmente, esse tempo é denominado tempo de espera.

A função densidade de probabilidade da distribuição exponencial envolve um parâmetro que é a constante positiva λ , cujo valor determina o local e a forma da função densidade (MUN, 2005).

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (7)$$

Já a função de distribuição cumulativa de X é:

$$F(x) = P(X \leq x) = \begin{cases} 1 - e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (8)$$

Segundo Mun (2005), a média e a variância de uma variável aleatória exponencial pode ser calculada usando integração por partes. O resultado está a seguir:

$$\mu_X = \frac{1}{\lambda} \quad (9)$$

$$\sigma_X^2 = \frac{1}{\lambda^2} \quad (10)$$

2.2.4 Distribuição Gama

A distribuição gama é uma distribuição contínua cuja finalidade é estender a utilidade da distribuição exponencial no modelamento de tempos de espera. Ela envolve uma determinada integral conhecida como função gama (NAVIDI, 2012).

Segundo Mun (2006), uma variável aleatória que é distribuída sob gama com parâmetro de forma k e escala θ é denotada por $X \sim \Gamma(k, \theta) \equiv \text{Gama}(k, \theta)$. A função densidade de probabilidade é definida por:

$$f(x, k, \theta) = \frac{x^{k-1} e^{-x/\theta}}{\theta^k \Gamma(k)} \quad \text{para } x > 0 \text{ e } k, \theta > 0 \quad (11)$$

A função de distribuição acumulada é definida por:

$$F(x, k, \theta) = e^{-x/\theta} \sum_{i=k}^{\infty} \frac{1}{i!} \left(\frac{x}{\theta}\right)^i \quad \text{para } x > 0 \text{ e } k, \theta > 0 \quad (12)$$

A média e a variância nesta distribuição são definidas por:

$$\mu_X = k\theta \quad (13)$$

$$\sigma_X^2 = k\theta^2 \quad (14)$$

2.2.5 Distribuição Beta

A distribuição Beta é muito flexível e é comumente usada para representar a variabilidade de um intervalo fixo. Uma das aplicações mais importantes da distribuição Beta é o seu uso como uma distribuição conjugada para o parâmetro da distribuição de Bernoulli. Nesta aplicação, a distribuição Beta é usada para representar a incerteza na probabilidade de ocorrência de um evento.

A distribuição Beta é frequentemente usada para modelar a proporção pertencente ao intervalo $(0,1)$, pois essa distribuição está definida neste intervalo. A função densidade de probabilidade Beta pode ser definida por:

$$f(x, \alpha, \beta) = \frac{x^{\alpha-1} (1-x)^{\beta-1}}{B(\alpha, \beta)} \quad \text{onde } B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)} \quad (15)$$

A função de distribuição acumulada é definida por:

$$F(x, \alpha, \beta) = I_x \quad \text{onde } I_x \text{ é a função Beta incompleta regularizada} \quad (16)$$

Para modelar uma variável aleatória Y restrita a um intervalo finito (m, n) , basta aplicar a relação $Y = (n - m)X + m$.

A distribuição Beta de uma variável aleatória X é representada como:

$X \sim \text{Beta}(\alpha, \beta)$ (MUN, 2006).

A média e a variância de uma variável aleatória X com distribuição Beta é definida por:

$$\mu_X = \frac{\alpha}{\alpha + \beta} \quad (17)$$

$$\sigma_X^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} \quad (18)$$

2.3 Gerenciamento de risco

2.3.1 Riscos e Incertezas

Segundo Lo e Mueller (2010), as distinções entre os vários tipos de incertezas diferenciam claramente a economia da física. Os economistas conhecem essas distinções há várias décadas. Em 1921, o economista Frank Knight definiu dois tipos de aleatoriedade: um que é passível de análise estatística, que Knight chamou de "risco", e outro que não é passível de análise estatística, que ele chamou de "incerteza".

Lo e Mueller (2010), propuseram uma classificação mais refinada sobre incertezas, dividindo o tema em cinco níveis.

- Nível 1: Total Certeza - Este é o reino da física clássica, um mundo determinístico governado pelas leis de Newton. Todos os estados passados e

futuros do sistema são determinados exatamente se as condições iniciais são fixas e conhecidas, ou seja, nada é incerto. Mesmo assim, sabe-se que pequenas alterações nas condições iniciais podem resultar em grandes mudanças dos resultados futuros e o controle destas condições iniciais nem sempre é trivial.

- **Nível 2: Risco sem Incerteza** - Este nível de aleatoriedade foi o que Frank Knight classificou como “risco” em sua tese de doutorado pela Universidade de Chicago em 1921. Neste nível, as aleatoriedades existentes são governadas por uma conhecida distribuição de probabilidade de um conjunto completo de resultados conhecidos. No nível “Risco sem Incerteza”, a teoria da probabilidade e os modernos teoremas e modelos são extremamente úteis para análise de risco. Neste universo, nenhuma inferência estatística é necessária porque conhecemos exatamente as distribuições de probabilidades. Embora o resultado de uma determinada aposta seja desconhecido, sabe-se todas as regras e probabilidades. Nenhuma outra informação relevante para o resultado é ocultada. Esta seria a situação em um hipotético cassino honesto, onde as regras são claramente definidas e seguidas. Evidentemente, esta situação tem pouca similaridade com o mercado financeiro.
- **Nível 3: Redução Total das Incertezas** - Este é o risco com um grau de incerteza. Incerteza esta devido ao não conhecimento das probabilidades para um conjunto de resultados que presumimos serem completamente conhecidos. Neste nível, a inferência estatística clássica deve ser aplicada à teoria da probabilidade. Com o termo “Redução Total das Incertezas”, estamos nos referindo a situações onde a aleatoriedade pode ser aproximada à incerteza do Nível 2, desde de que se possua grande quantidade de dados e sejam utilizadas todas as ferramentas de análise estatística e de teoria de probabilidade. A “Redução Total das Incertezas” é muito semelhante a um cassino honesto, mas em que as probabilidades não são registradas, desta forma, as probabilidades devem ser deduzidas. Em termos gerais, a “Total Redução das Incertezas” descreve um mundo no qual um modelo único gera todos os resultados. Esse modelo é definido por um número finito de parâmetros desconhecidos que não mudam com o tempo e que podem ser estimados.

- **Nível 4: Redução Parcial das Incertezas** – Este nível de incerteza refere-se às situações onde existe um limite para o que se pode deduzir sobre os fenômenos que geram os dados. Embora as leis de probabilidade ainda operem nesse nível, existe um complexo grau de incerteza com relação a geração dos dados que não podem ser reduzidos ao nível 2 de incerteza, mesmo com uma quantidade infinita de dados. No nível de “Redução Parcial das Incertezas”, estamos em um cassino que pode ou não ser honesto, e as regras tendem a mudar de tempos em tempos sem aviso prévio. Nessa situação, a estatística clássica pode não ser tão útil quanto uma perspectiva bayesiana.
- **Nível 5: Incerteza Irredutível** – Este nível trata do total desconhecimento que não se pode reduzir com uma maior coleta de dados ou com a utilização de métodos de inferência estatística mais sofisticados ou ainda com a utilização de supercomputadores. Esse tipo de incerteza é o domínio de filósofos e líderes religiosos, que se concentram não apenas no desconhecido, mas no que não se pode conhecer utilizando a razão.

Lo e Mueller (2010), ainda tratam do conceito de “Incerteza do Continuum” onde as condições e níveis de incertezas variam ao longo do tempo à medida que as tecnologias avançam e a compreensão dos fenômenos se aprofunda. Por exemplo, 3.000 anos atrás eclipses solares eram presságios misteriosos que teriam sido considerados Nível 5 de incerteza, mas hoje esses eventos são bem compreendidos e podem ser previstos com total certeza (nível 1).

2.3.2 A Natureza do Risco

A pesquisa inovadora de Harry Markowitz, ganhador do Prêmio Nobel, sobre a natureza do risco e retorno, revolucionou o mundo das finanças. Seu trabalho que hoje é conhecido em todo o mundo como a Fronteira Eficiente de Markowitz, olha para a natureza do risco e do retorno. Markowitz não olhou para o risco como um inimigo, mas como uma condição que deve ser abraçada e equilibrada através de seus retornos esperados. Posteriormente, o conceito de risco e retorno foi aperfeiçoado por William Sharpe. Em seu trabalho, William Sharpe desenvolveu o modelo de precificação de ativos de capital (CAPM), onde a taxa de retorno exigida de um ativo de risco é equivalente ao retorno de um ativo livre de risco somado a uma medida de

risco beta sistemática e não-diversificada, multiplicada por um prêmio de retorno do risco de mercado. Em resumo, um ativo de maior risco requer um maior retorno. No modelo de Markowitz pode-se encontrar um equilíbrio entre risco e retorno, onde dependendo da propensão ao risco de um investidor os melhores retornos podem ser obtidos através da fronteira eficiente (MUN, 2010).

2.3.3 A Estatística do Risco

O estudo das estatísticas refere-se à: coleta, apresentação, análise e utilização de dados numéricos, para inferir e tomar decisões diante de incertezas onde os dados reais da população são desconhecidos. Existem dois tipos de estudos de estatística: a estatística descritiva, onde os dados são resumidos, e a estatística inferencial, onde a população é generalizada através de uma pequena amostra aleatória de forma que, a amostra se torne útil para se fazer previsões ou tomar decisões quando as características da população são desconhecidas.

Os parâmetros calculados da amostra (por exemplo: média, mediana, desvio padrão) são denominados estatísticas, esses parâmetros sugerem que toda a população foi pesquisada e os resultados são tabulados. Assim, na tomada de decisão, a estatística é muito importante. Muitas vezes a população inteira é ainda desconhecida, visto que seria muito demorado ou extremamente custoso o total conhecimento da população (MUN, 2010).

Usualmente, na estatística inferencial, realizam-se as seguintes etapas:

- Projetar o experimento. Esta fase inclui a criação de maneiras de coletar todos os dados possíveis e relevantes;
- Coleta de dados da amostra. Os dados são coletados e tabulados;
- Análise de dados. A análise estatística é realizada;
- Estimação ou previsão. As inferências são feitas com base nas estatísticas obtidas;
- Teste de hipóteses. As decisões são testadas em relação aos dados para ver os resultados;

- Qualidade de ajuste. Os dados reais são comparados aos dados históricos para avaliar quão precisa, válida e confiável é a inferência;
- Tomada de decisão. As decisões são tomadas com base no resultado da inferência.

2.3.3.1 Medindo o Centro da Distribuição – 1º Momento

Segundo Mun (2010), o primeiro momento de uma distribuição mede a taxa esperada de retorno de um projeto em particular. Mede a localização dos cenários do projeto e possíveis resultados em média. As estatísticas comuns para o primeiro momento incluem: média, mediana (centro de uma distribuição) e moda (valor de maior ocorrência). A Figura 01 ilustra o primeiro momento onde, neste caso, o primeiro momento dessa distribuição é medido pela média (μ) ou valor médio.

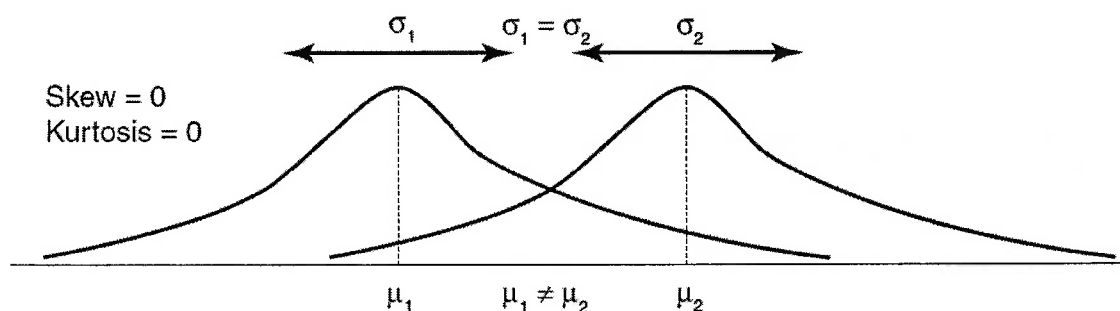


Figura 01 - Primeiro momento. Fonte: Mun (2010).

Primeiro momento: média aritmética ou média simples (comumente usados).

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (19)$$

2.3.3.2 Medindo o *Spread* da Distribuição – 2º Momento

O segundo momento mede a dispersão de uma distribuição, que é uma medida de risco. O *spread* ou largura de uma distribuição, mede o quanto a variável se afasta da média. Uma pequena dispersão indica que os dados são mais concentrados próximos da média, isto é, menor risco. Já uma distribuição com grande dispersão indica que os dados se afastam da média, ou seja, maior risco. A Figura 02

ilustra duas distribuições com primeiros momentos idênticos (médias idênticas), mas segundos momentos (riscos) muito diferentes (MUN, 2010).

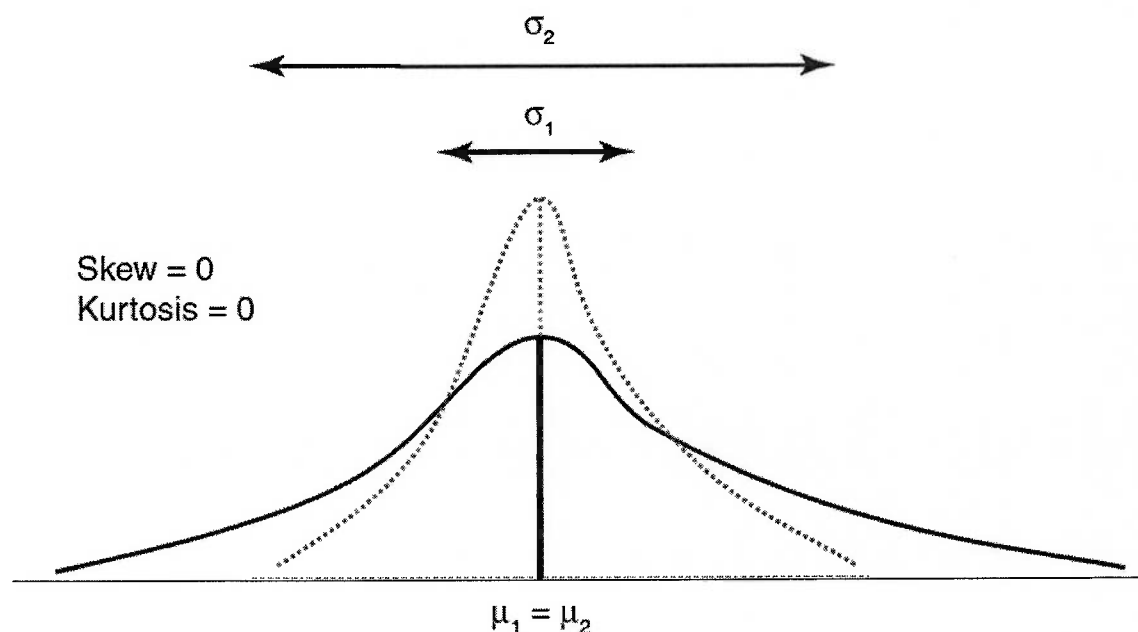


Figura 02 - Segundo momento. Fonte: Mun (2010).

A visualização fica mais clara na Figura 03, onde é apresentado um exemplo do desempenho de duas ações. Os movimentos da primeira ação (ilustrados pela linha pontilhada) possuem menor flutuação. Já os movimentos da segunda ação (ilustrado pela linha mais escura) possuem flutuação de preço com *spread* muito maior. Entretanto, ambas possuem o mesmo valor médio (primeiro momento) (MUN, 2010).

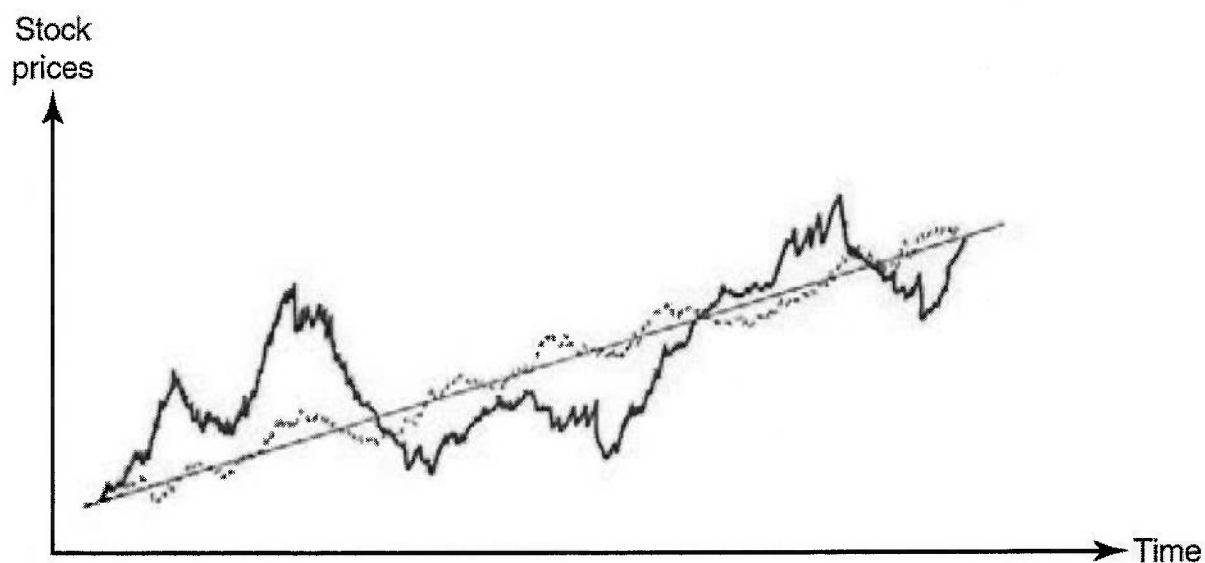


Figura 03 - Flutuação de preços de duas ações com mesma média. Fonte: Mun (2010).

Segundo momento: Desvio padrão (comumente usado).

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \quad (20)$$

2.3.3.3 Medindo a Inclinação da Distribuição – 3º Momento

Segundo Mun (2010), o terceiro momento mede a assimetria de uma distribuição, isto é, como a distribuição é voltada para um lado ou para o outro. A Figura 04 ilustra uma assimetria negativa ou inclinação para esquerda (a cauda dos pontos de distribuição aponta para a esquerda). A Figura 05 ilustra uma assimetria positiva ou inclinação para direita (a cauda da distribuição aponta para a direita). (MUN, 2010).

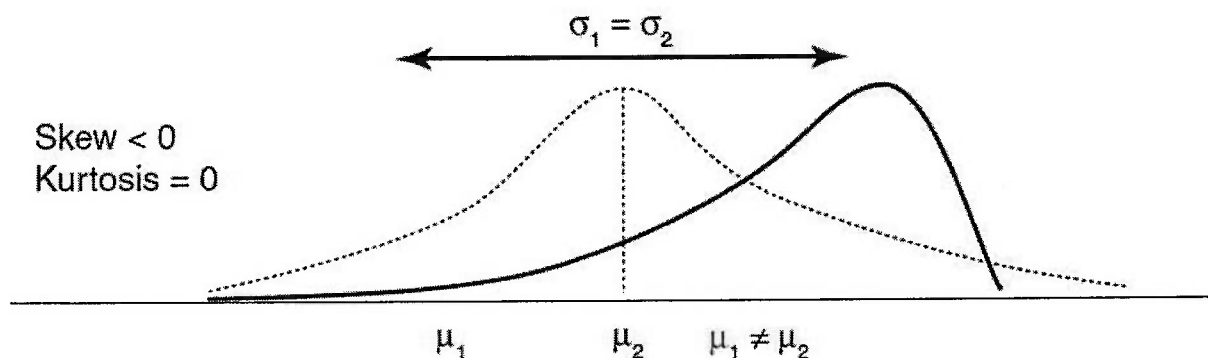


Figura 04 - Terceiro momento – assimetria negativa. Fonte: Mun (2010).

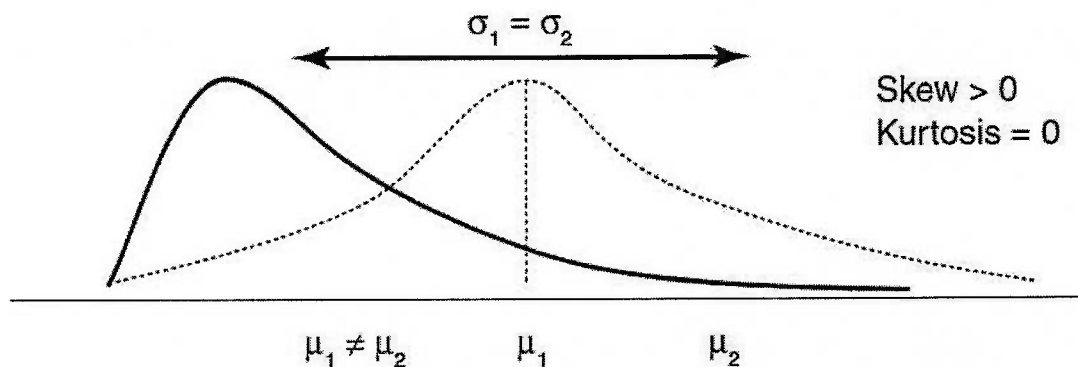


Figura 05 - Terceiro momento – assimetria positiva. Fonte: Mun (2010).

$$assimetria = \frac{n}{(n-1)(n-2)} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma} \right)^3 \quad (21)$$

2.3.3.4 Medindo os Eventos de Cauda Catastróficos – 4º Momento

O quarto momento, ou curtose, mede o pico de uma distribuição. A Figura 06 ilustra esse efeito, onde a linha pontilhada representa uma distribuição normal com um excesso de curtose de 0. Por sua vez, a linha contínua representa uma distribuição que possui uma curtose maior, onde a área sob a curva é mais espessa nas caudas com menor área no corpo central, esta condição gera grandes influências na análise de risco. Na distribuição de maior curtose, os três primeiros momentos (média, desvio padrão e assimetria) podem ser idênticos, mas o quarto momento (curtose) será diferente. Esta condição significa que, embora os retornos e os riscos sejam idênticos, as probabilidades de eventos extremos e catastróficos (grandes perdas potenciais ou grandes ganhos) ocorrerem são maiores para uma distribuição de curtose. Desta forma, ignorar a curtose de uma distribuição pode ser muito prejudicial para a análise (MUN, 2010).

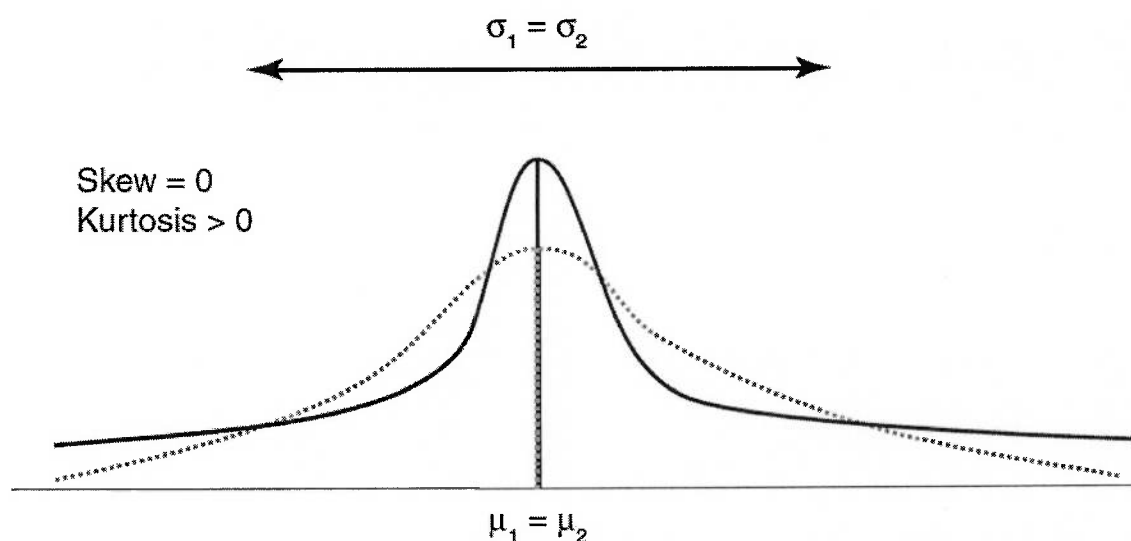


Figura 06 - Quarto momento – Curtose. Fonte: Mun (2010).

$$curtose = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma} \right)^4 - \frac{3(n-1)^2}{(n-2)(n-3)} \quad (22)$$

2.3.4 Medindo Risco

Segundo Mun (2010), existem várias maneiras de medir o risco em projetos. Abaixo serão apresentadas as medidas de risco mais comuns, bem como seus potenciais benefícios e falhas.

- a) **Probabilidade de Ocorrência:** Essa abordagem é simplista e ainda eficaz. Por exemplo, considere que após uma análise de risco de um projeto, constatou-se que existe uma probabilidade de 10% deste projeto apresentar prejuízos nos próximos 5 anos. Esta abordagem claramente fornece ao investidor uma mensuração do risco envolvido. Uma outra forma de exemplificar a utilidade desta abordagem é apresentada supondo dois projetos semelhantes, com custos de implantação e retornos esperados idênticos. Neste exemplo, se a análise de risco for desprezada, a escolha de qual projeto implantar é indiferente. Entretanto, considere que uma análise de risco é executada em ambos os projetos. O primeiro projeto apresenta uma probabilidade de 70% de

ocorrer perdas e o segundo projeto apresenta uma probabilidade de apenas 5%. Claramente o segundo projeto é melhor, quando os riscos são analisados;

- b) **Desvio padrão e variância.** O desvio padrão é uma mensuração da média do desvio de cada ponto dos dados em relação à média. O desvio padrão é uma medida de risco usual, quanto maior for seu valor maior será sua distribuição e, portanto, tem um risco maior. A desvantagem desta medida é que ambas as variações de alta e baixa são incluídas no cálculo do desvio padrão.
- c) **Desvio Semi-Padrão.** O desvio semi-padrão mede apenas o desvio padrão dos riscos negativos e ignora o lado positivo das flutuações. Modificações do desvio semi-padrão incluem o cálculo somente dos valores abaixo da média ou valores abaixo de um limite (por exemplo, lucros negativos ou fluxos de caixa negativos). Isso proporciona uma melhor ideia do risco negativo;
- d) **Volatilidade.** O conceito de volatilidade é amplamente utilizado nas aplicações de opções reais e pode ser definido brevemente como uma medida de incertezas e riscos. A volatilidade pode ser estimada usando vários métodos numéricos, tais como: EWMA e GARCH. A volatilidade é mais difícil de definir e estimar. Entretanto, esta é uma medida de risco mais sofisticada.
- e) **Beta.** Beta é uma medida de risco decorrente da exposição a movimentos gerais de mercado em oposição a fatores idiossincráticos. Um beta abaixo de 1 pode indicar um investimento com menor volatilidade financeira do que o mercado, ou um investimento volátil cujos movimentos de preço não são altamente correlacionados com o mercado. Um beta maior que 1 geralmente significa que o ativo é volátil e tende a subir e descer com o mercado. Ele não mede o risco de um investimento mantido em uma base individual, mas a quantidade de risco que o investimento acrescenta a uma carteira já diversificada.
- f) **Coeficiente de variação.** O coeficiente de variação é simplesmente definido como a razão do desvio padrão pela média, onde os riscos são comuns. Por exemplo, a distribuição de um grupo de alturas de alunos (medido em metros) pode ser comparado com a distribuição dos pesos dos alunos (medidos em quilogramas). Esta medida de risco ou dispersão é aplicável quando as estimativas, as medidas, as magnitudes ou as unidades diferem entre si;
- g) **Value at Risk.** *Value at Risk* (VaR) ficou famoso por J. P. Morgan na década de 1990 através da introdução da sua abordagem *Risk Metrics*, a qual tem sido

utilizada por várias instituições financeiras ao redor do mundo. O *Value at Risk* é um método para avaliar o risco em operações financeiras, o qual resume em um número, o risco de um produto financeiro ou o risco de uma carteira de investimentos de um montante financeiro. Este número representa a pior perda esperada em um dado horizonte de tempo e é associado a um intervalo de confiança.

- h) **Pior cenário ou arrependimento.** Outra medida de risco bastante simples é o valor do pior cenário dadas as perdas catastróficas. Nesta situação, após uma avaliação financeira do projeto avalia-se a probabilidade de ocorrência deste evento catastrófico e seu impacto na tomada de decisão do projeto.
- i) **Retorno ajustado ao risco do capital.** Retorno ajustado ao risco do capital (RAROC) considera a razão da diferença entre o percentil 50 do retorno (mediana) e o retorno do quinto percentil de um projeto pelo desvio padrão. Essa abordagem é usada principalmente pelas instituições financeiras para estimar retornos sujeito a riscos, medindo apenas a potencial desvantagem e ignorando o balanço positivo.

2.3.5 Simulação de Monte Carlo para Análise de Risco

Segundo Mun (2010), a simulação de Monte Carlo cria cenários futuros gerando milhares e até centenas de milhares de exemplos de caminhos de resultados e analisa suas características predominantes. Na prática, o método numérico de Monte Carlo pode ser usado para análise, quantificação e prevenção de riscos. Uma alternativa à simulação de Monte Carlo é o uso de modelos estocásticos complexos de forma fechada. Ainda segundo Mun (2010), quando modelado corretamente, o método de Monte Carlo fornece respostas semelhantes aos métodos matematicamente mais complexos. Além disso, existem muitas aplicações na vida real onde os modelos de forma fechada não possuem recursos disponíveis para simular o comportamento do fenômeno.

A simulação de Monte Carlo, em resumo, é um número aleatório útil para previsão, estimativa e análise de risco. Uma única simulação calcula vários cenários de um modelo escolhendo valores de forma repetida de acordo com uma distribuição de probabilidade definida pelo usuário. Como todos os cenários produzem resultados associados a um modelo, cada cenário pode ter uma previsão.

A simulação de Monte Carlo é um tipo de simulação paramétrica, onde parâmetros específicos de distribuição são necessários para que a simulação seja realizada.

2.4 Utilização do método de Monte Carlo no mercado imobiliário

No mercado imobiliário o uso do método de Monte Carlo é bastante limitado, supostamente devido à sua demanda por compreensão matemática e estatística (HOESLI; JANI; & BENDER, 2006). Embora tenha seu uso limitado, alguns acadêmicos fizeram algumas importantes pesquisas sobre o tema. Stephen Pyhrr introduziu um modelo de simulação para medir o risco de um investimento imobiliário. A pesquisa combinou a análise de investimentos imobiliários tradicionais com premissas probabilísticas nas variáveis de entrada do modelo FCD. Grande parte das pesquisas posteriores sobre o tema basearam-se na pesquisa de Pyhrr (PYHRR, 1973). Hoesli et al. usaram o método de Monte Carlo para prever valores de propriedade com taxa de desconto variando durante o período de investimento (HOESLI; JANI; & BENDER, 2006). French et al. utilizaram o método de Monte Carlo para analisar as incertezas relacionadas à previsão de investimentos (FRENCH & GABRIELLI, 2005).

Suhonen (2014), em sua dissertação de mestrado, apresentou a utilização do método de Monte Carlo para realizar a avaliação de um shopping center que havia sido vendido para uma empresa do setor. O modelo proposto por Suhonen tratou as taxas de desocupação e os preços dos aluguéis do shopping como variáveis aleatórias que seguiam distribuições de probabilidade. O trabalho foi comparado com um modelo de FCD utilizado para a avaliação e negociação do empreendimento. O resultado desta comparação revelou a fragilidade do método FCD em tratar os riscos e incertezas desta avaliação.

Crudden (2012), apresenta o Método de Monte Carlo como uma importante ferramenta incremental ao método FCD para avaliação de empreendimentos imobiliários comerciais. Crudden (2012), apresenta ainda uma lista interessante dos fatores de risco em que o método de Monte Carlo pode ser empregado para auxiliar o método FCD.

Risco	Comentários
Despesas de capital	Trata da incerteza na previsão dos gastos de capital necessários para concluir uma obra física. Gastos extras podem resultar em um efeito inverso nos futuros fluxos de caixa.
Taxa de capitalização	Trata da incerteza da taxa utilizada para obter o financiamento necessário para a construção do empreendimento. O movimento da taxa de capitalização tem uma relação inversa com o valor da propriedade.
Risco de construção	Trata da incerteza na conclusão da atividade de construção, em relação ao tempo e ao custo previsto.
Taxa de desconto	Trata da incerteza de qual taxa utilizar para descontar os fluxos de caixas futuros.
Período de desocupação	Trata da incerteza em prever com precisão o período de desocupação de um imóvel de locação.
Taxa de aluguel	Trata da incerteza de prever o valor que o mercado estará disposto a pagar pelo aluguel do imóvel.
Taxa de crescimento do aluguel	Trata da incerteza da taxa de crescimento ou decrescimento do valor do aluguel no futuro.
Taxa de inadimplência	Trata da incerteza relacionada a taxa de inadimplência dos imóveis alugados.

Infelizmente, não foram encontrados estudos envolvendo a utilização do Método de Monte Carlo para simular a atividade de *flip* no mercado de *real estate* norte americano.

3. Desenvolvimento do Modelo

3.1 Aprofundamento da caracterização do objeto de estudo

O objetivo deste trabalho é desenvolver uma modelagem estocástica utilizando método de Monte Carlo para avaliar o investimento na atividade de flip no mercado de *real estate* norte americano. Para isso faz-se necessário a realização de um procedimento que pode ser descrito em cinco etapas.

Na primeira etapa, deve-se conhecer em detalhes o investimento, suas formas de geração de caixa, sua estrutura financeira, seus ativos e passivos, ou seja, necessita-se de todo o conhecimento para a realização de um modelo determinístico. Quanto maior for o conhecimento obtido do funcionamento desta atividade financeira, melhor e mais preciso será o modelo. É necessário identificar de forma específica a composição dos fluxos de caixa deste projeto, seus padrões e variâncias.

A segunda etapa trata da identificação e classificação das variáveis com forte influência na modelagem do problema, ou seja, variáveis com maior significância para o modelo. A partir de então, deve-se coletar, com base na etapa anterior, uma amostra de dados adequada da atividade de investimento e a realização do correto tratamento matemático a ser utilizado em cada variável, seja ela uma variável estocástica ou uma variável determinística.

A terceira etapa consiste na determinação das distribuições de probabilidade das variáveis estocásticas determinadas na segunda etapa. Neste processo, deve-se verificar a distribuição de probabilidade que melhor representa cada variável. Neste momento, a distribuição de probabilidade de uma variável estocástica será adaptada considerando a amostra coletada. Além disso, um tratamento estatístico pode ser necessário para o tratamento de *outliers*.

A quarta etapa consiste na determinação de correlações entre as diversas variáveis do modelo. Todas as significativas correlações entre as variáveis devem ser consideradas para a geração dos diversos cenários da avaliação. As variáveis aleatórias que serão geradas na modelagem devem obedecer às correlações encontradas.

Na quinta e final etapa do modelo ocorre a geração dos milhares de possíveis cenários que seguem as características da amostra (correlação entre variáveis, distribuições das variáveis e interação entre elas) que diferentemente de uma abordagem padrão (FCD, múltiplos, cenários, etc...), apresentam uma distribuição de

probabilidade contendo um mapeamento dos diversos possíveis resultados do investimento em análise, apresentando uma visão mais realística do problema além de garantir a possibilidade de uma análise estatística do risco.

3.2. Abordagem do objeto do estudo

Serão abordados nesta sessão a metodologia e objetivos para o desenvolvimento de um modelo numérico, utilizando simulação de Monte Carlo, para avaliação da viabilidade financeira da atividade de flip no mercado de *real estate* norte americano.

O objetivo de uma atividade de flip consiste na aquisição de um imóvel, sua renovação ou reforma e por fim, sua comercialização com obtenção de lucros. A atividade de investimento no mercado americano de *real estate* possui uma série de fatores de risco de mercado, mesmo este sendo muito mais estável e confiável que o mercado imobiliário brasileiro. Assim sendo, neste trabalho será apresentada uma abordagem matemática para a modelagem deste problema, utilizando o método de Monte Carlo, visando gerenciar de forma mais competente os riscos inerentes desta atividade. Vale ressaltar que este trabalho trata apenas dos riscos de mercado, riscos de crédito não serão abordados.

Pustejovsky (2017), afirma que a atividade de flip no mercado de *real estate* norte americano vem se tornando cada vez mais popular, criando grande concorrência para o setor. Além disso, atualmente ocorre a diminuição da oferta de residências confiscadas pelos bancos. Assim sendo, neste trabalho, pretende-se avaliar se esta atividade de investimento ainda é economicamente interessante e quais são os riscos envolvidos nesta atividade.

3.3 Modelagem

Neste modelo serão avaliados apenas imóveis residenciais do tipo: casas unifamiliares (maior incidência no mercado residencial americano), dois e três dormitórios de classe média. Devido à grande variação e flutuação dos valores de compra e venda em função da localização, serão consideradas apenas propriedades dos condados de Essex, Union e Middlesex do estado de Nova Jersey, as quais fazem parte da região metropolitana de Nova York. Os 3 condados acima descritos contam com uma população de aproximadamente 2.215.000 de habitantes.

Serão consideradas para este modelo a existência de 5 variáveis (todas expressas em dólares americanos), sendo elas:

- X_1 : Preço de compra da casa;
- X_2 : Gastos com despesas operacionais, legais e inspeções para aquisição das casas;
- X_3 : Gastos com renovação e reforma das residências;
- X_4 : Valor de venda da casa;
- X_5 : Lucro ou prejuízo na operação (P&L).

Existiriam ainda as variáveis tempo e taxa de desconto que seriam usadas para o cálculo do valor presente, entretanto como o ciclo de uma atividade de *flip* é usualmente de 3 a 4 meses e ainda considerando que a taxa livre de risco dos estados unidos é baixa, essas variáveis foram desprezadas neste modelo.

As variáveis X_1 , X_2 , X_3 e X_4 serão tratadas como variáveis aleatórias que irão compor a variável final X_5 (P&L). As variáveis X_1 , X_2 , X_3 e X_4 serão simuladas conforme suas distribuições de probabilidade e suas correlações entre si. O P&L da operação será dado pela relação:

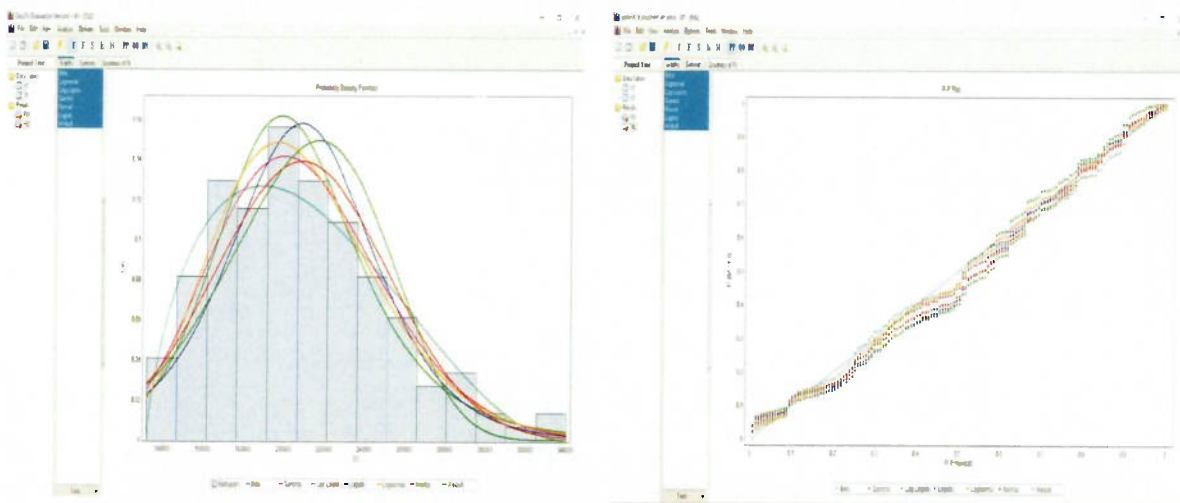
$$X_5 = X_4 - (X_1 + X_2 + X_3) \quad (23)$$

A coleta de dados será realizada pelo acesso ao banco de dados MLS (Multiple Listing Service). O MLS trata de um grande banco de dados inicialmente criado em 1907 pelos corretores de valores imobiliários norte-americanos que necessitavam de auxílio para venda de seus imóveis (WIKIPEDIA THE FREE ENCYCLOPEDIA, 2008). Desta forma, toda nova movimentação imobiliária de cada imóvel é compartilhada neste

banco de dados. Este banco de dados possibilita o conhecimento do histórico de compra e venda de propriedades, datas de realização de negociações além de informações financeiras e de financiamentos. O acesso a este banco de dados ocorre apenas por meio de corretores imobiliários registrados. Neste trabalho, o acesso a este banco de dados será realizado com o auxílio de um escritório de *real estate* do condado de Union que concordou em cooperar com o fornecimento dos dados com o intuito exclusivo de uso dos dados para fins acadêmicos. Não serão fornecidos dados de caráter pessoal (nomes, SSN ou endereços dos imóveis).

De posse da base de dados das variáveis X_1 , X_2 , X_3 e X_4 será realizada a adaptação (fit) das variáveis nas funções de distribuição de probabilidade que tenham melhor aderência a amostragem. Optou-se por usar apenas distribuições de probabilidade contínuas devido à amostra ser pequena.

Para auxiliar na determinação da função de distribuição que mais se adapta a cada caso será utilizado o software EasyFit que realiza o ajuste dos dados a diversas distribuições de probabilidade selecionadas. Além disso, o software também realiza testes de hipótese e aderência.



a) Função de densidade de probabilidade ajustada a base de dados.

b) P-P Plot

Figura 07 - Software EasyFit.

Neste passo, serão realizadas as adaptações das distribuições de probabilidade (δ) nas variáveis aleatórias X_2 , X_3 e X_4 . As distribuições que visualmente não apresentarem nenhuma relação com a variável X_i serão automaticamente descartadas. Posteriormente, serão realizados testes de hipóteses em cada uma das distribuições selecionadas. Cada teste consistirá em:

- Hipótese nula $H_0 \rightarrow$ Não rejeita H_0 quando: $X_i \in \delta$; Rejeita H_a
- Hipótese nula $H_a \rightarrow$ Não rejeita H_a quando: $X_i \notin \delta$; Rejeita H_0
- Serão considerados os testes de aderência de Kolmogorov-Smirnov, Anderson-Darling e Chi-Quadrado, os quais considerarão a hipótese nula H_0 não rejeitada, ou seja, $X_i \in \delta$ se o valor estatístico encontrado pelo respectivo teste for menor que o valor de significância de 5%.

A distribuição de probabilidade que apresentar melhor resultado nos testes de aderência será escolhida para descrever a variável em análise X_i .

Para o tratamento dos *outliers* será adotada a análise da distância entre quartis por meio de gráficos de Box-plot.

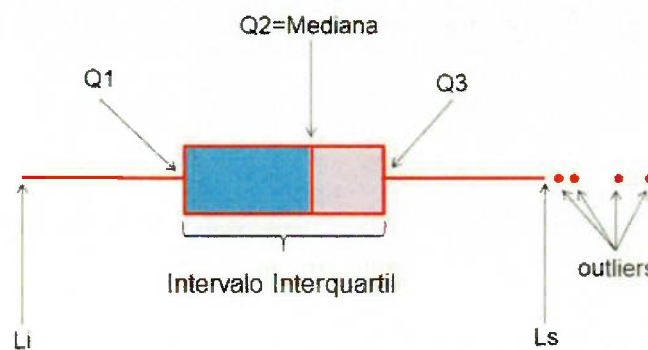


Figura 08 - Box-plot ou diagrama de caixa.

Onde:

Q1: primeiro quartil;

Q3: terceiro quartil;

Li: Limite inferior $\rightarrow \max\{\min(\text{dados}; Q1 - 1.5(Q3 - Q1))\}$;

Ls: Limite superior $\rightarrow \min\{\max(\text{dados}; Q3 - 1.5(Q3 - Q1))\}$;

Intervalo Interquartil = $Q3 - Q1$.

Neste trabalho optou-se pela eliminação dos *outliers* que forem encontrados nos dados.

Com a identificação e eliminação dos *outliers* da amostra e com as funções de distribuição de probabilidade e seus parâmetros definidos, inicia-se a simulação de Monte Carlo com a geração dos N cenários que seguem suas respectivas distribuições de probabilidade. Posteriormente, realiza-se a padronização de X_i , por:

$\left(\frac{x_i - \bar{x}_i}{\sigma_i}\right)$. Aplica-se então a fatoração de Cholesky para considerar a correlação entre as variáveis X_i . A fatoração de Cholesky consiste na determinação de uma matriz $M \mid M.M^T = \Sigma$. Seja ϕ as variáveis aleatórias correlacionadas tal que $\phi = X.M$ (WILMOTT, 2006).

Consequentemente, retorna-se à condição inicial de X_i , realizando: $(\bar{X}_i + \phi_i, \sigma_i)$.

Por fim, calcula-se o P&L pela equação (23) e procede-se para a avaliação dos resultados obtidos.

3.4 Obtenção dos resultados

De posse do modelo numericamente desenvolvido, conforme metodologia apresentada anteriormente, parte-se para a obtenção, exploração e análise da variável de saída X_5 , ou seja, do lucro ou prejuízo na atividade de *flip* no mercado de *real estate* norte americano. Uma das saídas do modelo desenvolvido será o histograma de densidade de probabilidade dos lucros e prejuízos deste investimento. A Figura 10 apresenta uma ideia deste resultado.

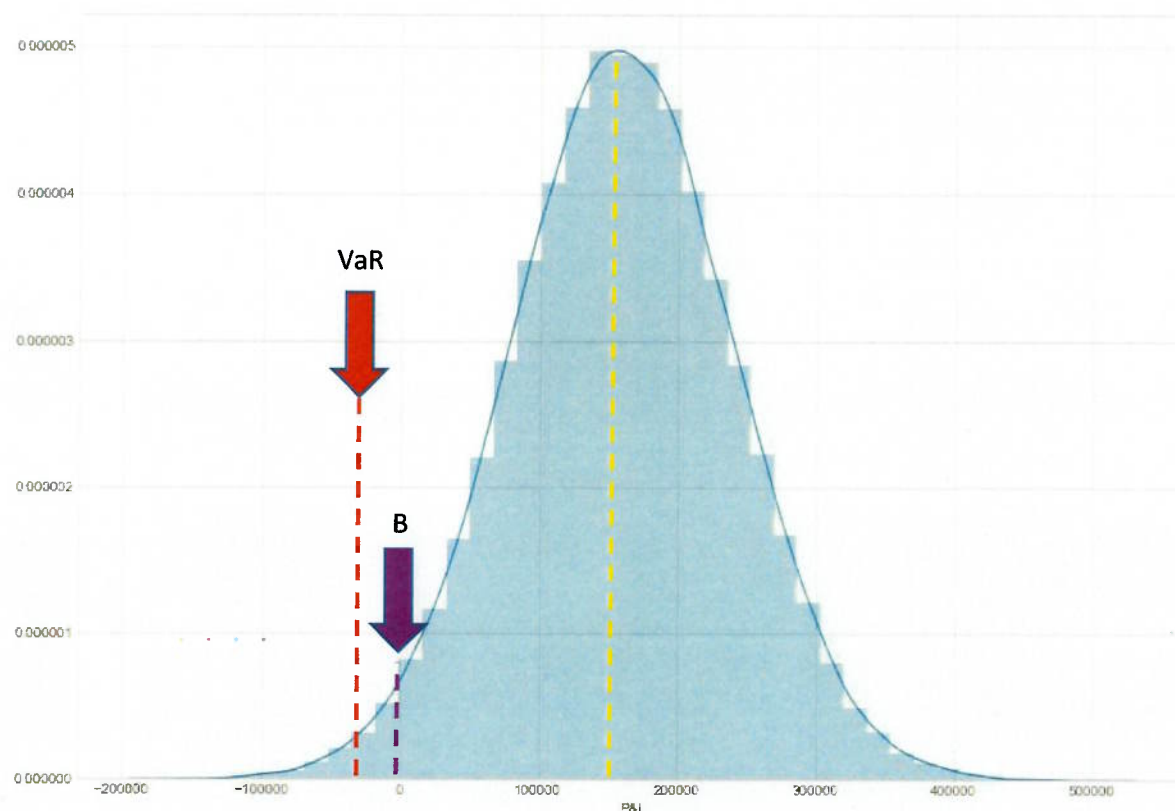


Figura 09 - Distribuição de probabilidade hipotética do P&L do investimento.

Observando a Figura 09 apresentada acima, inicialmente podemos verificar a forma da distribuição dos resultados, o que será extremamente interessante para verificar o comportamento da variável. Podemos também encontrar o ponto de Break-Even, indicado pela seta roxa da Figura 09, ou seja, o ponto onde não ocorre ganhos nem prejuízos. Neste caso hipotético o Break-Even encontrado foi no percentil 2,34%, assim, em condições normais de mercado, existe a probabilidade de 97,66% de o investidor realizar lucro nesta operação. Outra medida interessante de avaliação de risco que pode ser obtida é o valor propenso ao risco, ou VaR (*Value at Risk*), neste caso considerando um VaR de 1% obteve-se o valor de -\$24.193,82, ou seja, em condições normais de mercado, o investidor está aceitando tomar um risco neste investimento com 1% de probabilidade de ter um prejuízo de \$24.193,82 ou mais.

Pode-se também calcular o valor esperado para o retorno deste investimento, que neste caso foi de \$159.886,00.

Além disso, outro resultado interessante que pode ser obtido do modelo é a relação do P&L com as demais variáveis (X_1 , X_2 , X_3 e X_4). Desta análise pode ser obtido, por exemplo, o comportamento dos cenários que resultaram em prejuízo. Quais seus padrões de comportamento? E como podem ser evitados ou minimizados? Desta mesma maneira, quais são os casos em que os lucros foram acima da média? Quais as relações entre as variáveis nestes casos?

Resultados como estes auxiliarão o investidor na minimização do risco e na tomada de decisão.

4. Aplicação da Teoria a um Caso Específico

4.1 Coleta e tratamento estatístico da amostra

Foi realizada a coleta de dados de uma população de casas residenciais unifamiliares, de padrão médio, com 2 e 3 dormitórios, com algum tipo de manutenção necessária, no período entre 2016 a 2018, na região dos condados de Essex, Union e Middlesex do estado de Nova Jersey - EUA. Nesta amostra foram coletadas inicialmente 436 unidades, ou seja, 436 negociações de compra e venda, entretanto esta base de dados continha informações faltantes. As variáveis: gastos com despesas operacionais (X_2) e gastos com renovação e reforma das residências (X_3) foram mais difíceis de obter pois, o banco de dados MLS não possui estas informações e para sua obtenção necessitou-se entrar em contato com proprietários ou empresas que prestaram serviços de reforma e renovação nestas unidades. Assim sendo, a amostra inicial foi reduzida, obtendo-se uma amostragem completa com 147 unidades. As Figuras 10, 11 e 12 apresentam os dados da amostra, suas correlações e possíveis funções de distribuição de probabilidade, respectivamente.

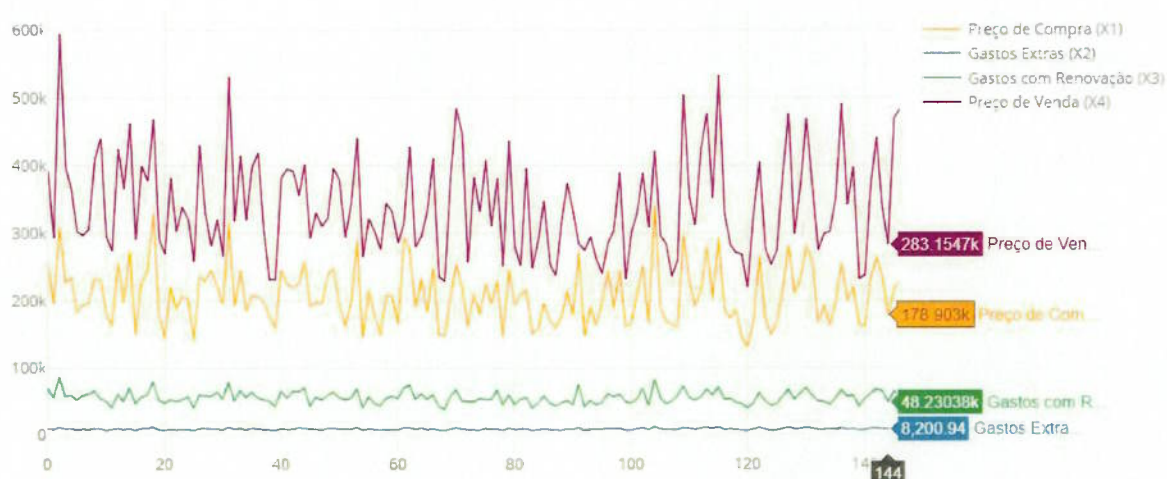


Figura 10 - Amostragem completa obtida das variáveis X_1 , X_2 , X_3 , X_4 .

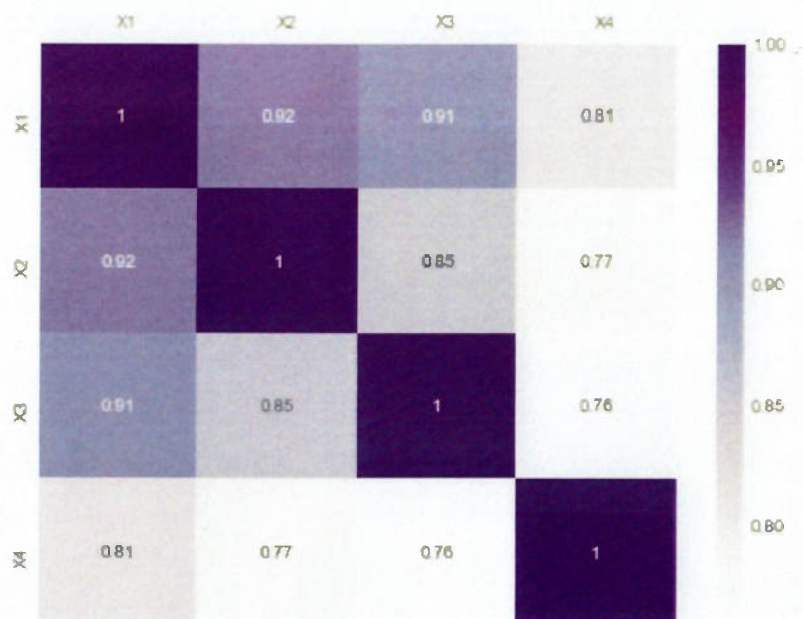


Figura 11 - Matriz de correlação das variáveis X_1 , X_2 , X_3 , X_4 .

Percebe-se uma forte correlação entre as variáveis, especialmente as correlações entre o preço de compra do imóvel com os gastos extras (X_1 - X_2) e com os gastos com renovação (X_1 - X_3).

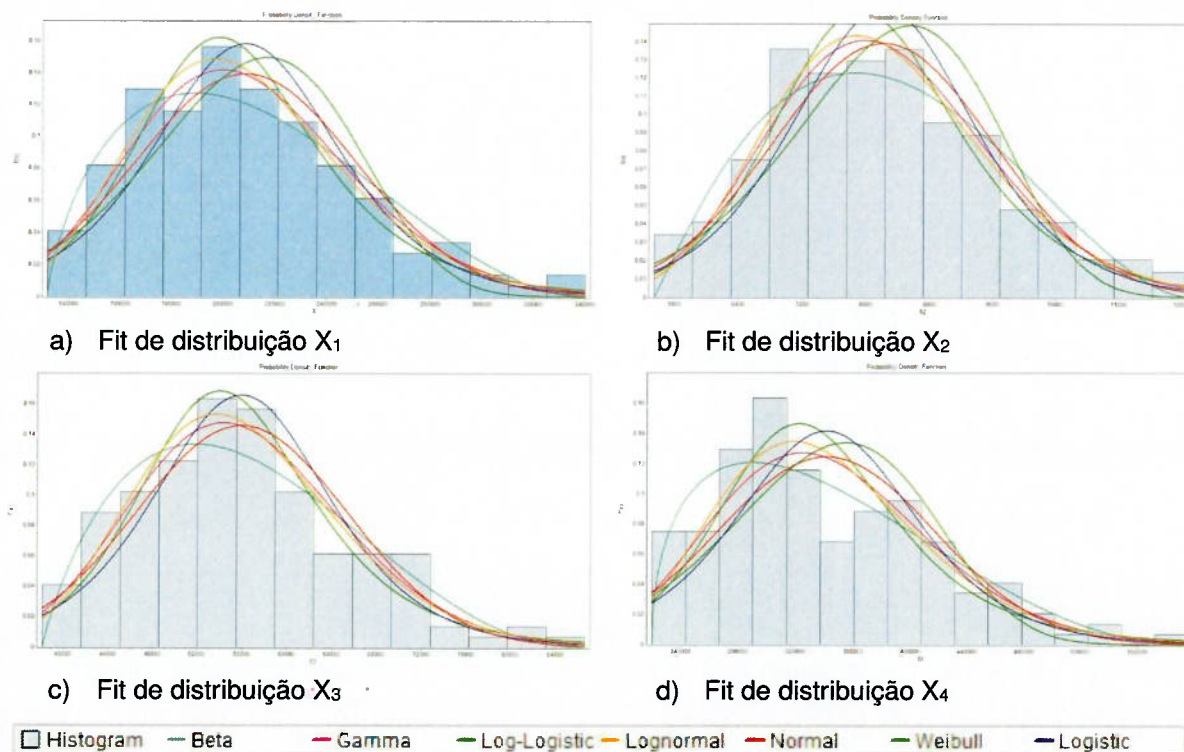


Figura 12 - Fit das funções de densidade de probabilidade aplicado as variáveis X_1 , X_2 , X_3 e X_4 .

Posteriormente, realizou-se o tratamento dos *outliers* com o uso dos gráficos de Box-Plot.

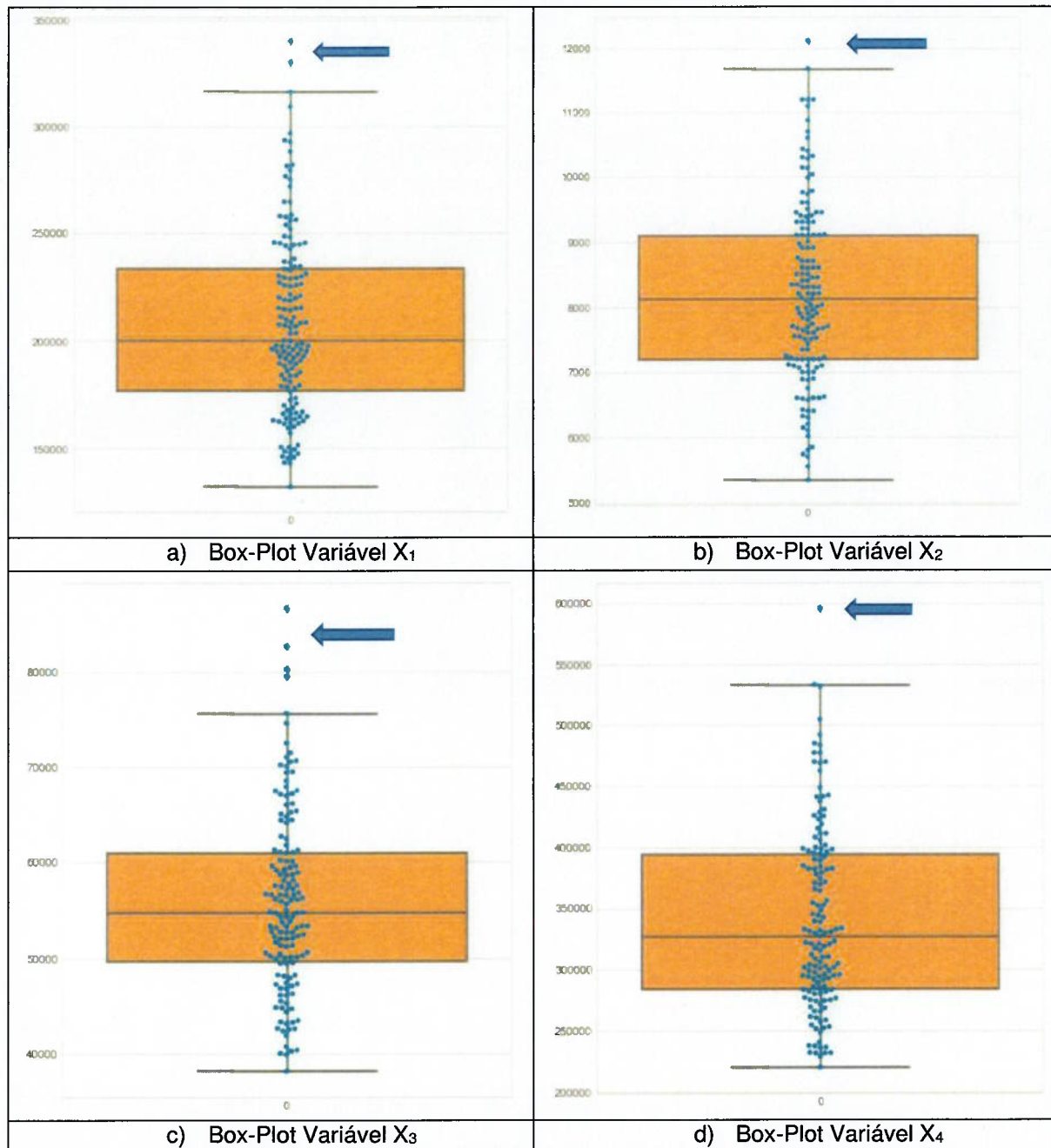


Figura 13 - Gráficos Box-Plot das variáveis X_1 , X_2 , X_3 e X_4 .

Na Figura 13, as setas azuis indicam a presença dos *outliers* que foram excluídos da amostra. No total foram eliminados 4 *outliers* da amostra, ficando a mesma com 143 exemplares.

4.2 Fit das distribuições de probabilidade do modelo

O próximo passo compreende a determinação das funções de densidade de probabilidade e suas respectivas propriedades.

4.2.1 Fit da distribuição de probabilidade da variável X_1

Abaixo são apresentadas as funções de densidade de probabilidade que visualmente melhor se adaptaram a variável X_1 . As funções que visualmente não tinham nenhuma relação de aderência com a variável X_1 foram diretamente descartadas.

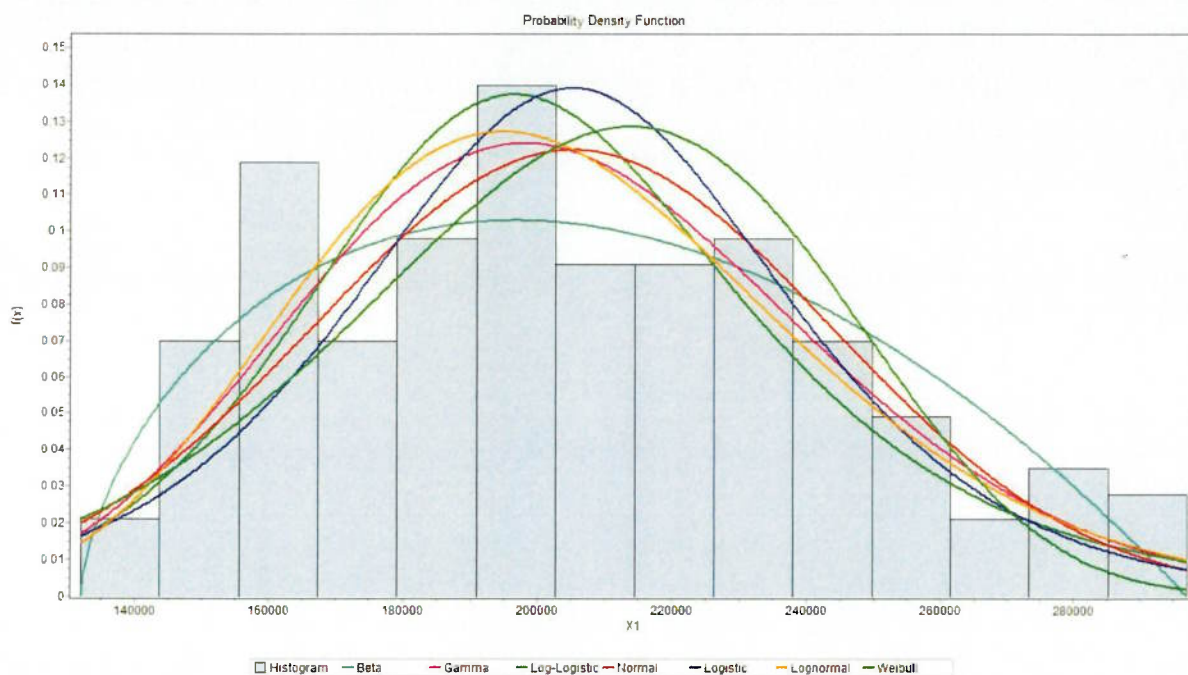


Figura 14 - Funções densidade de probabilidade ajustadas a variável X_1 sem outliers.

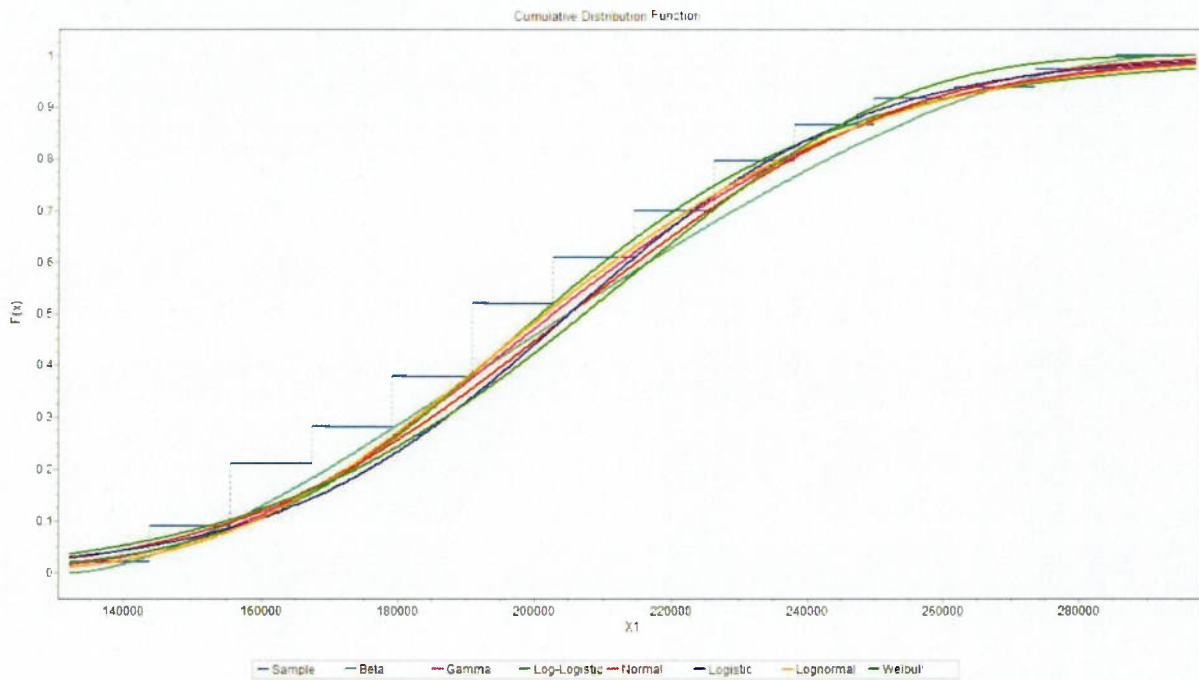


Figura 15 - Funções acumulada de probabilidade ajustadas a variável X_1 , sem outliers.

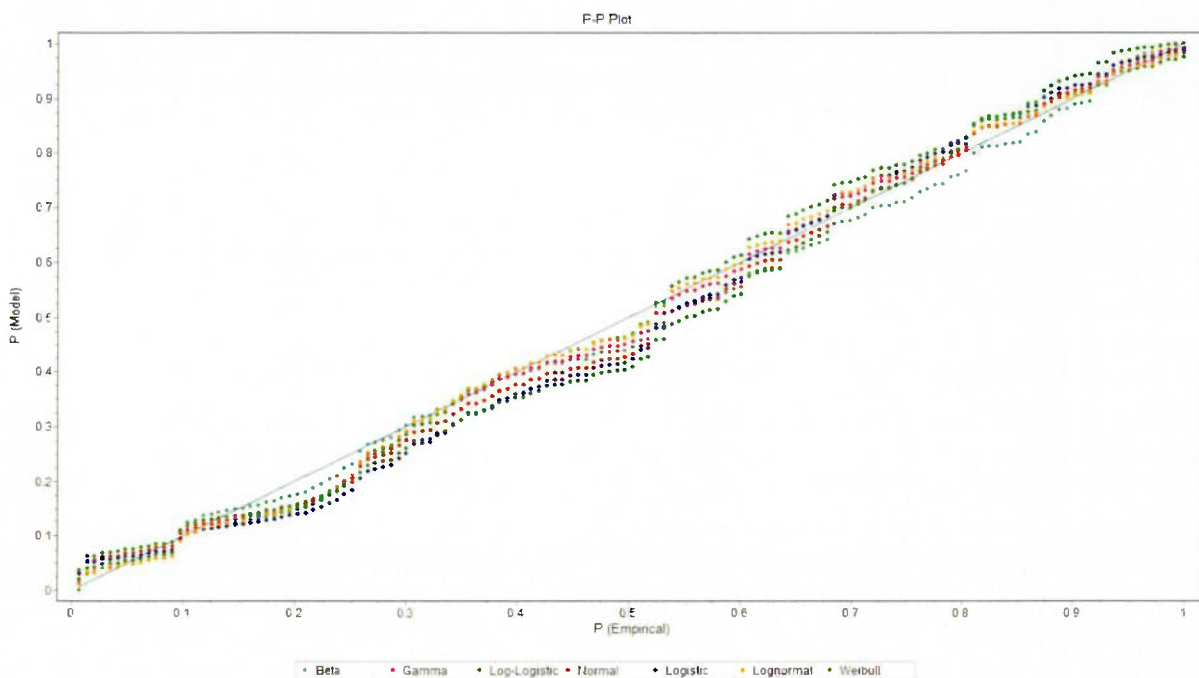


Figura 16 - PP-Plot da variável X_1 , sem outliers.

Posteriormente foram realizados os testes de hipótese e aderência para cada uma das 7 funções ajustadas à variável X_1 , considerando:

- Hipótese nula $H_0 \rightarrow$ Não rejeita H_0 quando: $X_i \in \delta$; Rejeita H_a
- Hipótese nula $H_a \rightarrow$ Não rejeita H_a quando: $X_i \notin \delta$; Rejeita H_0

O resumo dos testes de aderência para a variável X_1 é mostrado na Tabela 01. Em geral, obteve-se o melhor resultado com a distribuição gama:

Tabela 01 - Resumo dos testes de aderência variável X_1 .

Distribution	Kolmogorov Smirnov		Anderson Darling		Chi-Squared	
	Statistic	Rank	Statistic	Rank	Statistic	Rank
Gamma	0,0518	1	0,44629	1	3,8924	3
Lognormal	0,05445	2	0,46304	2	6,6424	4
Beta	0,05925	3	2,543	7	3,5175	2
Log-Logistic	0,0624	4	0,76806	3	10,237	7
Normal	0,07216	5	0,78513	4	1,8719	1
Logistic	0,08129	6	1,347	5	9,5758	6
Weibull	0,09585	7	2,0858	6	7,6848	5

A Tabela 02 apresenta o detalhamento dos resultados dos testes de hipótese para a variável X_1 utilizando a função de distribuição gama.

Tabela 02 - Detalhamento dos resultados dos testes de hipótese variável X_1 .

Kolmogorov-Smirnov		Anderson-Darling		Chi-Squared	
Sample Size	143	Sample Size	143	Deg. of freedom	7
Statistic	0,0518	Statistic	0,44629	Statistic	3,8924
Rank	1	Rank	1	Rank	3
α	5,00%	α	5,00%	α	5,00%
Critical Value	0,11356	Critical Value	2,5018	Critical Value	14,067
Reject H_0 ?	No	Reject H_0 ?	No	Reject H_0 ?	No

O teste de hipótese não rejeitou a hipótese nula para nenhum dos testes de aderência com um nível de significância de 5%. Assim sendo, a distribuição gama foi a escolhida para descrever a variável X_1 deste modelo com os parâmetros: $k(\text{forma})=28,515$; $\theta(\text{escalar})=7194,2$. Ou seja, $X_1 \sim \Gamma(28,515, 7194,2)$.

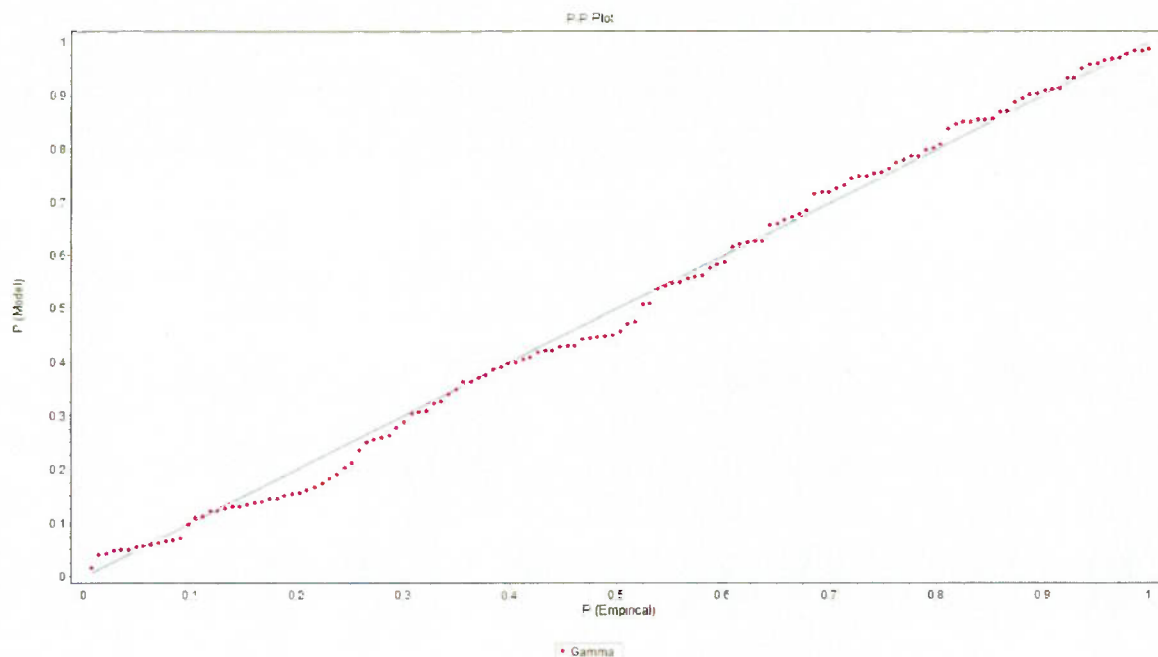


Figura 17 - PP-Plot da variável X_1 , somente com a distribuição gama.

4.2.2 Fit das distribuições de probabilidade da variável X_2

Abaixo são apresentadas as funções de densidade de probabilidade que visualmente melhor se adaptaram a variável X_2 . As funções que visualmente não tinham nenhuma relação de aderência com a variável X_2 foram diretamente descartadas.

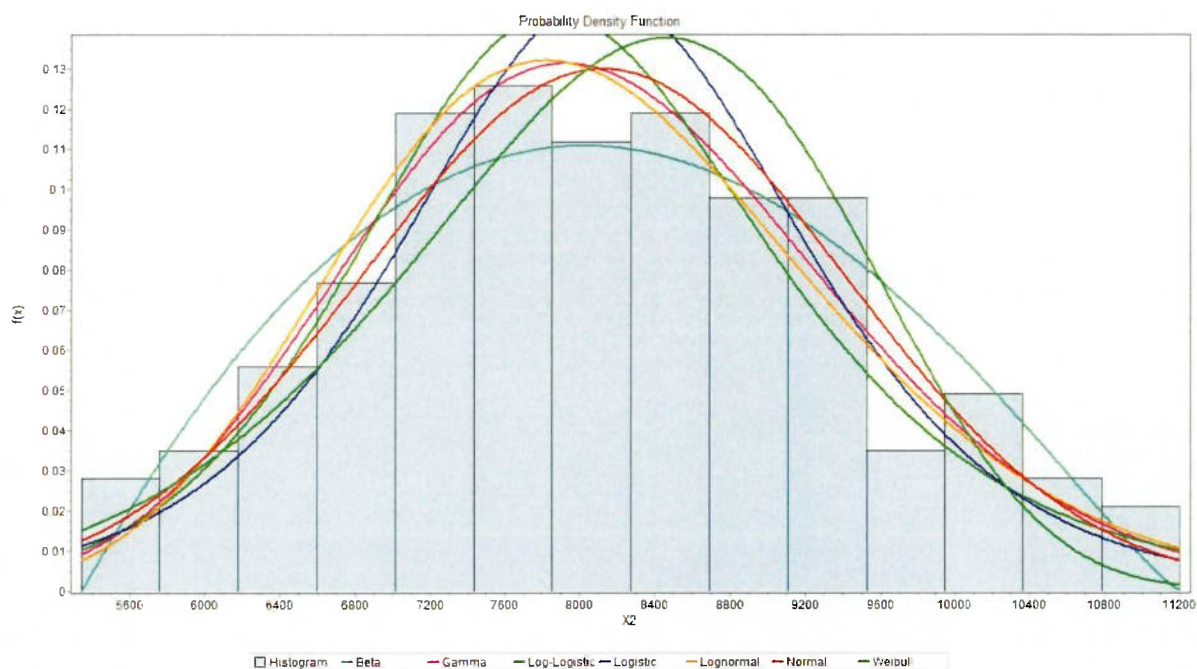


Figura 18 - Funções densidade de probabilidade ajustadas a variável X_2 sem outliers.

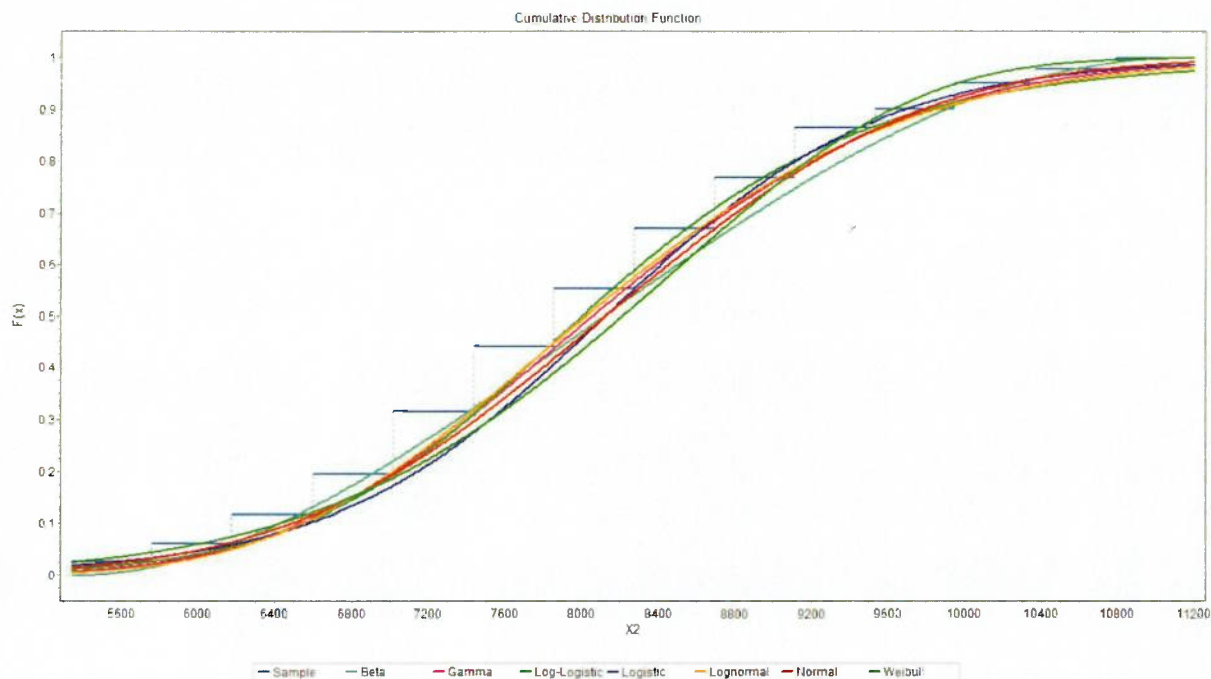


Figura 19 - Funções acumulada de probabilidade ajustadas a variável X_2 , sem outliers.

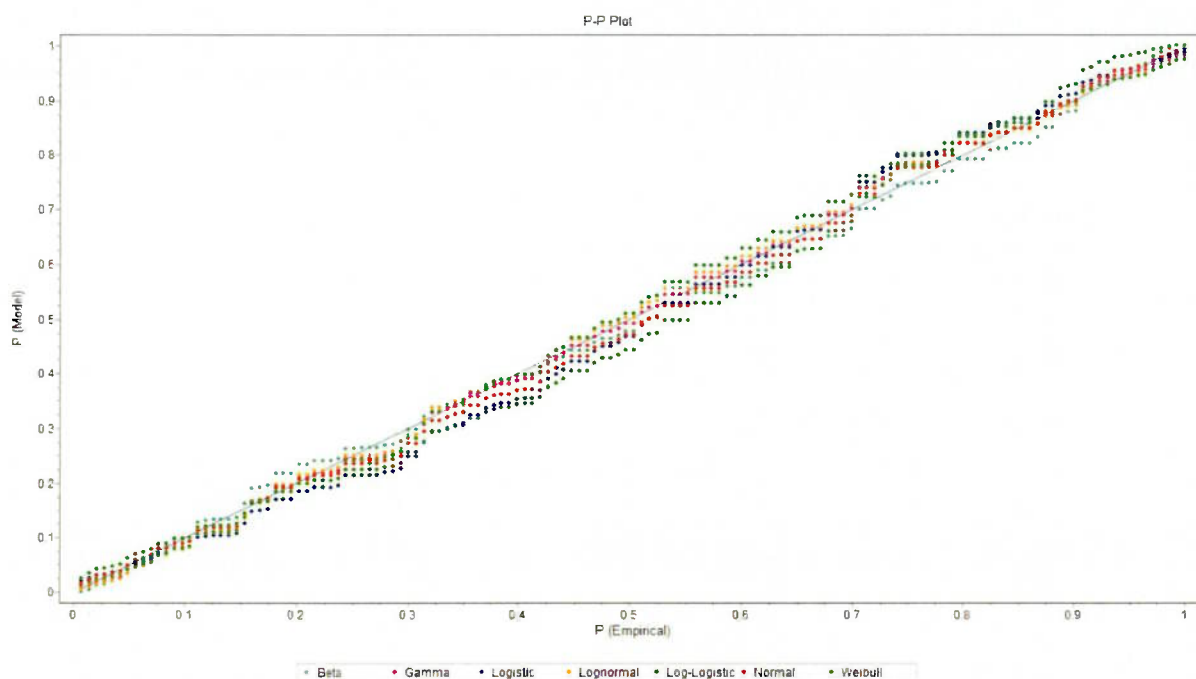


Figura 20 - PP-Plot da variável X_2 , sem outliers.

Posteriormente foram realizados os testes de hipótese e aderência para cada uma das 7 funções ajustadas à variável X_2 , considerando:

- Hipótese nula $H_0 \rightarrow$ Não rejeita H_0 quando: $X_i \in \delta$; Rejeita H_a
- Hipótese nula $H_a \rightarrow$ Não rejeita H_a quando: $X_i \notin \delta$; Rejeita H_0

O resumo dos testes de aderência para a variável X_2 é mostrado na Tabela 03. Em geral, obteve-se o melhor resultado com a distribuição normal:

Tabela 03 - Resumo dos testes de aderência variável X_2 .

Distribution	Kolmogorov Smirnov		Anderson Darling		Chi-Squared	
	Statistic	Rank	Statistic	Rank	Statistic	Rank
Normal	0,04476	1	0,26831	3	3,0428	2
Beta	0,04483	2	2,2874	7	1,7273	1
Gamma	0,0474	3	0,19485	1	3,7243	3
Lognormal	0,04931	4	0,25638	2	5,2517	5
Log-Logistic	0,06569	5	0,49934	4	3,9642	4
Weibull	0,06663	6	1,3588	6	5,9612	6
Logistic	0,06742	7	0,72091	5	5,9924	7

A Tabela 04 apresenta o detalhamento dos resultados dos testes de hipótese para a variável X_2 utilizando a função de distribuição normal.

Tabela 04 - Detalhamento dos resultados dos testes de hipótese variável X_2 .

Kolmogorov-Smirnov		Anderson-Darling		Chi-Squared	
Sample Size	143	Sample Size	143	Deg. of freedom	7
Statistic	0,04476	Statistic	0,26831	Statistic	3,0428
Rank	1	Rank	3	Rank	2
α	5,00%	α	5,00%	α	5,00%
Critical Value	0,11356	Critical Value	2,5018	Critical Value	14,067
Reject H_0 ?	No	Reject H_0 ?	No	Reject H_0 ?	No

O teste de hipótese não rejeitou a hipótese nula para nenhum dos testes de aderência com um nível de significância de 5%. Assim sendo, a distribuição normal foi a escolhida para descrever a variável X_2 deste modelo com os parâmetros: $\mu=8122,1$, $\sigma=1285,4$. Ou seja, $X_2 \sim N(8122,1, 1285,4)$.

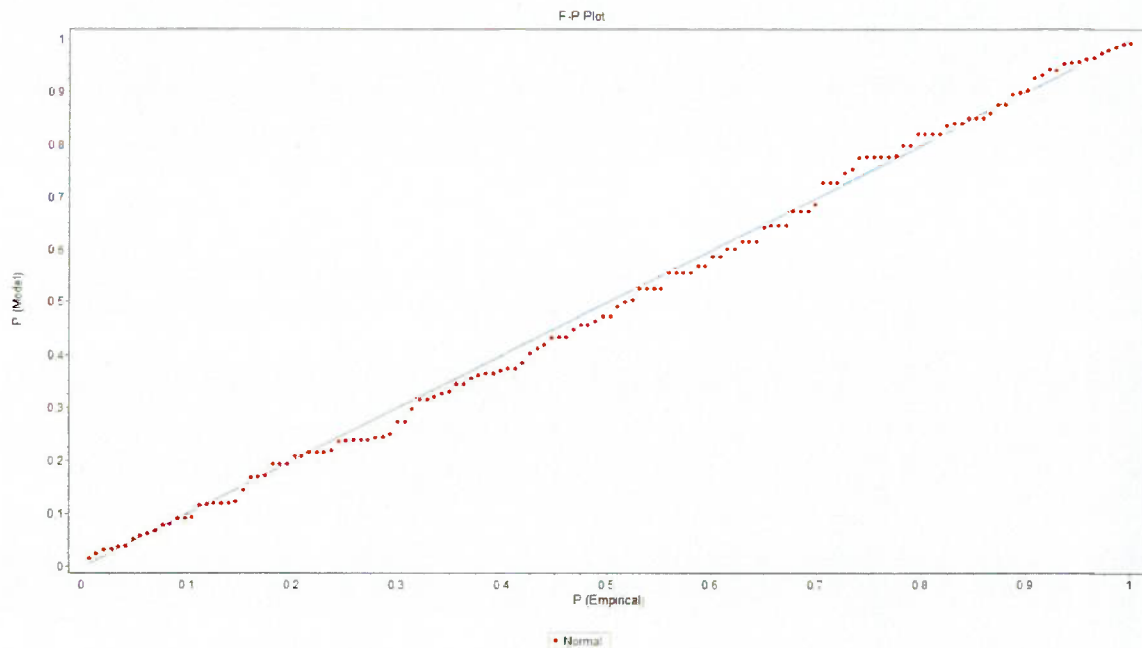


Figura 21 - PP-Plot da variável X_2 , somente com a distribuição normal.

4.2.3 Fit das distribuições de probabilidade da variável X_3

Abaixo são apresentadas as funções de densidade de probabilidade que visualmente melhor se adaptaram a variável X_3 . As funções que visualmente não tinham nenhuma relação de aderência com a variável X_3 foram diretamente descartadas.

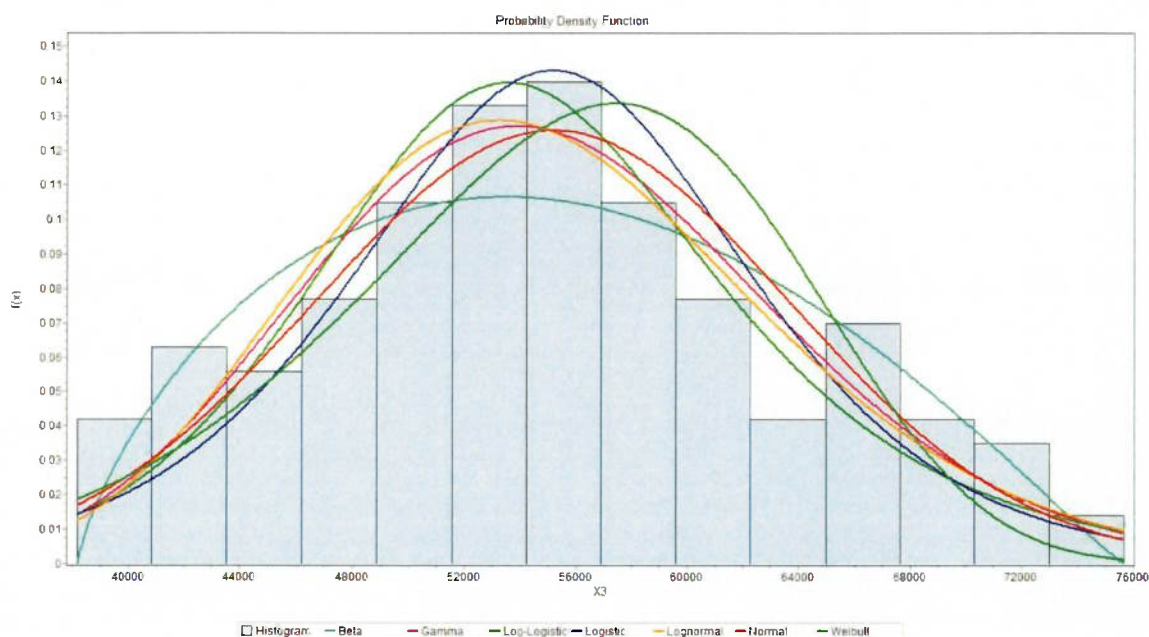


Figura 22 - Funções densidade de probabilidade ajustadas a variável X_3 sem outliers.

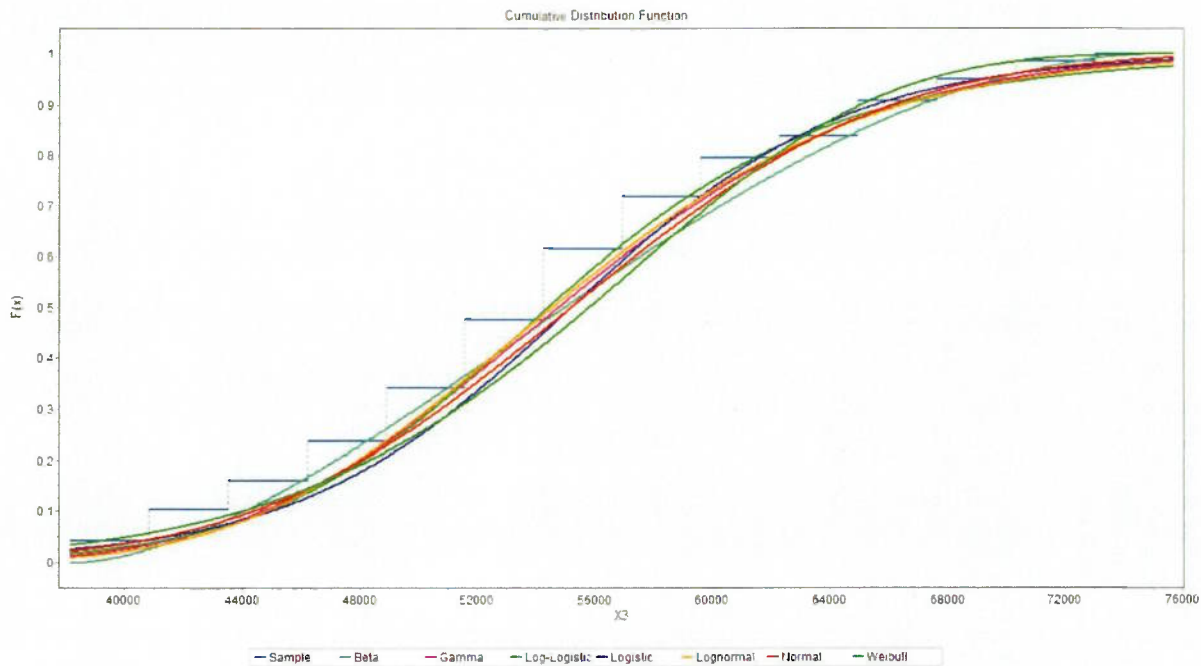


Figura 23 - Funções acumulada de probabilidade ajustadas a variável X_3 , sem outliers.

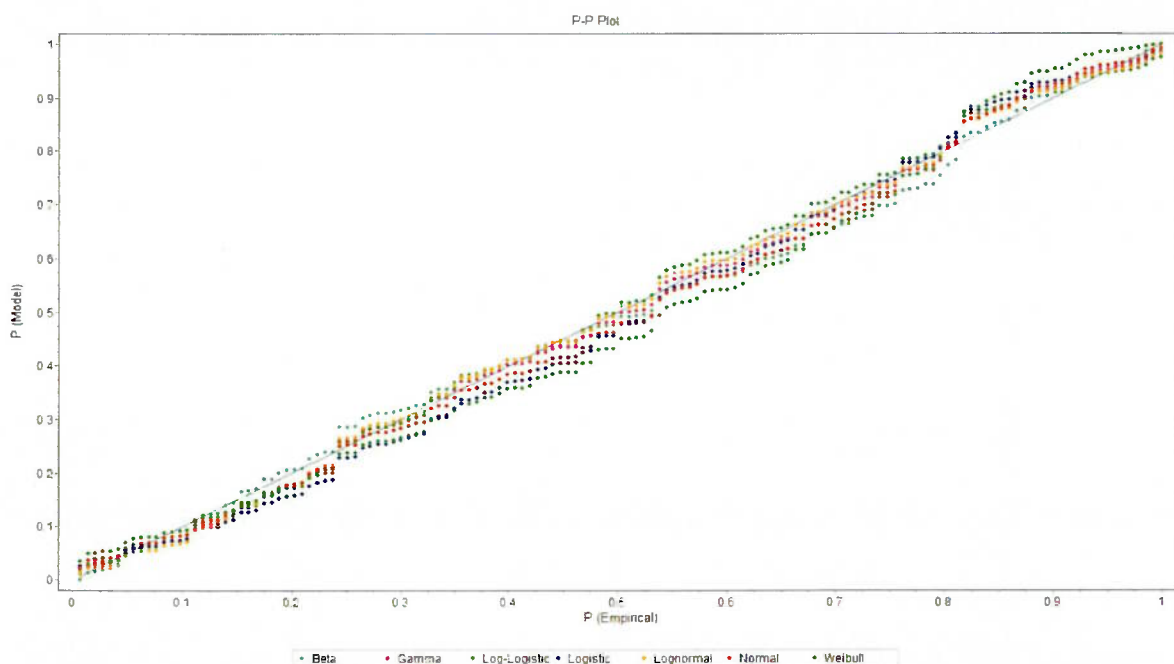


Figura 24 - PP-Plot da variável X_3 , sem outliers.

Posteriormente foram realizados os testes de hipótese e aderência para cada uma das 7 funções ajustadas à variável X_3 , considerando:

- Hipótese nula $H_0 \rightarrow$ Não rejeita H_0 quando: $X_i \in \delta$; Rejeita H_a
- Hipótese nula $H_a \rightarrow$ Não rejeita H_a quando: $X_i \notin \delta$; Rejeita H_0

O resumo dos testes de aderência para a variável X_3 é mostrado na Tabela 05. Em geral, obteve-se o melhor resultado com a distribuição gama:

Tabela 05 - Resumo dos testes de aderência variável X_3 .

Distribution	Kolmogorov Smirnov		Anderson Darling		Chi-Squared	
	Statistic	Rank	Statistic	Rank	Statistic	Rank
Gamma	0,04321	1	0,30587	1	5,0803	2
Lognormal	0,04322	2	0,32093	2	4,9402	1
Normal	0,04551	3	0,51495	4	5,8564	4
Log-Logistic	0,05336	4	0,48451	3	5,4492	3
Beta	0,05578	5	2,3702	7	7,8811	5
Logistic	0,06069	6	0,87665	5	9,9419	6
Weibull	0,0743	7	2,0457	6	14,575	7

A Tabela 06 apresenta o detalhamento dos resultados dos testes de hipótese para a variável X_3 utilizando a função de distribuição gama.

Tabela 06 - Detalhamento dos resultados dos testes de hipótese variável X_3 .

Kolmogorov-Smirnov		Anderson-Darling		Chi-Squared	
Sample Size	143	Sample Size	143	Deg. of freedom	7
Statistic	0,04321	Statistic	0,30587	Statistic	5,0803
Rank	1	Rank	1	Rank	2
α	5,00%	α	5,00%	α	5,00%
Critical Value	0,11356	Critical Value	2,5018	Critical Value	14,067
Reject H_0 ?	No	Reject H_0 ?	No	Reject H_0 ?	No

O teste de hipótese não rejeitou a hipótese nula para nenhum dos testes de aderência com um nível de significância de 5%. Assim sendo, a distribuição gama foi a escolhida para descrever a variável X_3 deste modelo com os parâmetros: $k(\text{forma})=42,169$; $\theta(\text{escalar})=1308,2$. Ou seja, $X_3 \sim \Gamma(42,169, 1308,2)$.

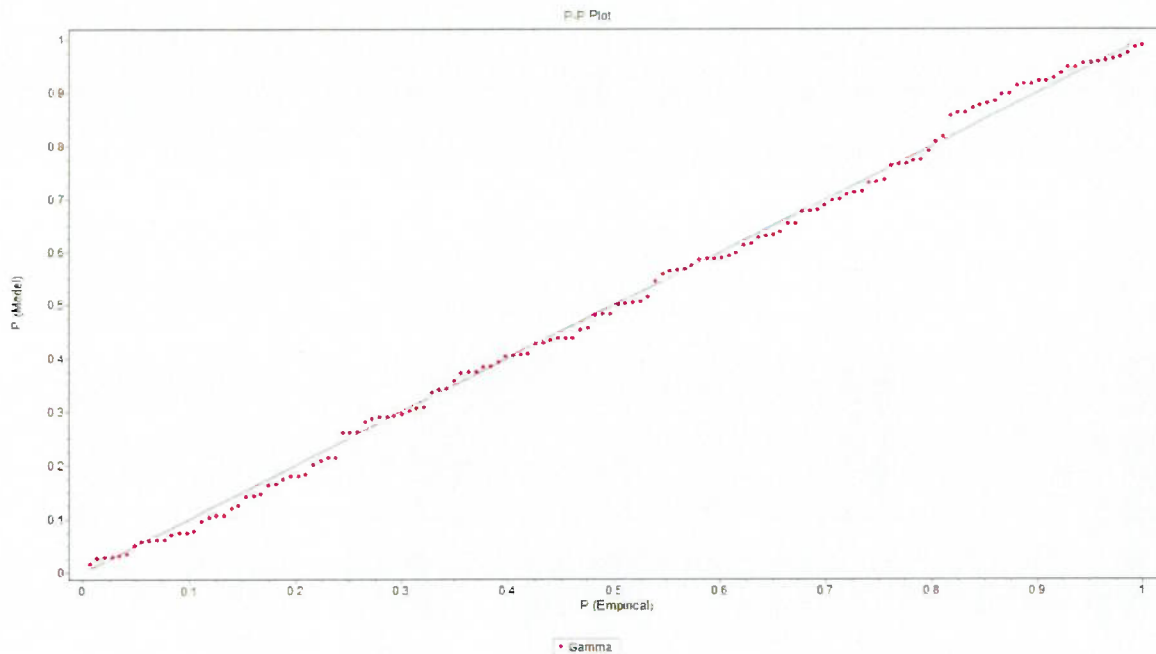


Figura 25 - PP-Plot da variável X_3 , somente com a distribuição gama.

4.2.4 Fit das distribuições de probabilidade da variável X_4

Abaixo são apresentadas as funções de densidade de probabilidade que visualmente melhor se adaptaram a variável X_4 . As funções que visualmente não tinham nenhuma relação de aderência com a variável X_4 foram diretamente descartadas.

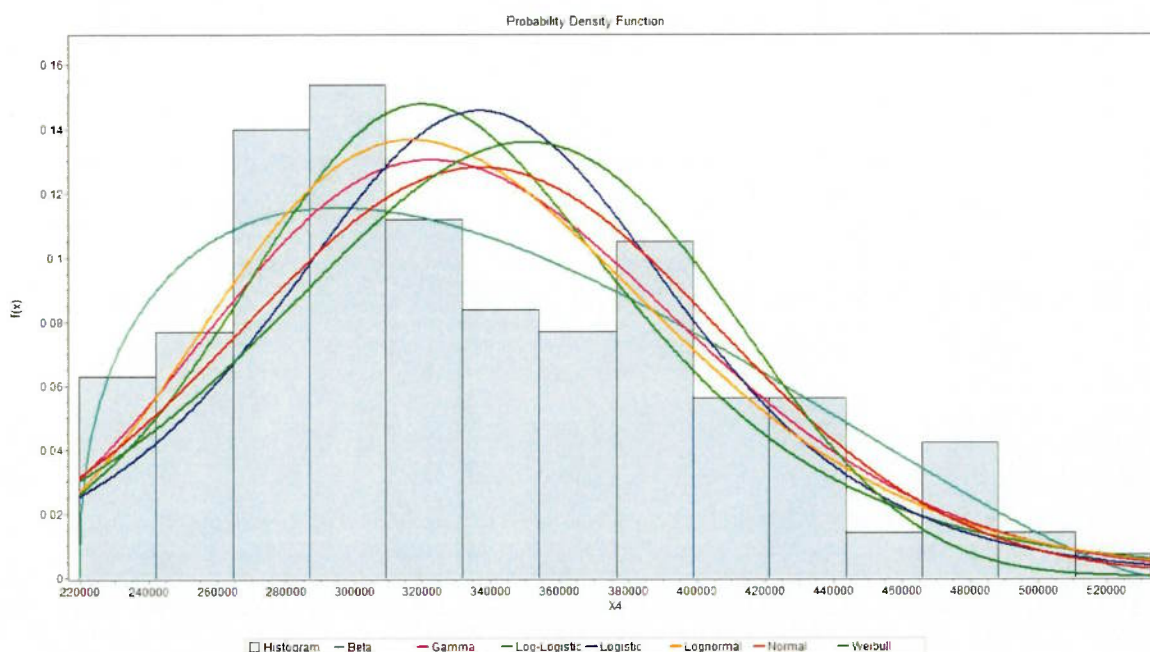


Figura 26 - Funções densidade de probabilidade ajustadas a variável X_4 sem *outliers*.

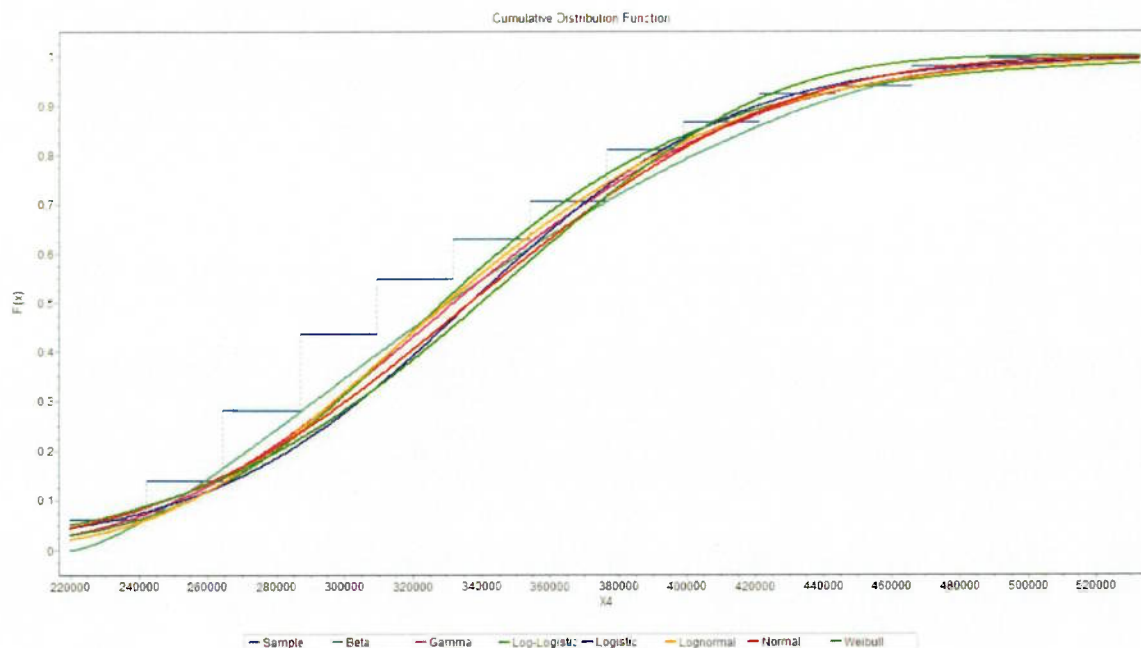


Figura 27 - Funções acumulada de probabilidade ajustadas a variável X_4 , sem outliers.

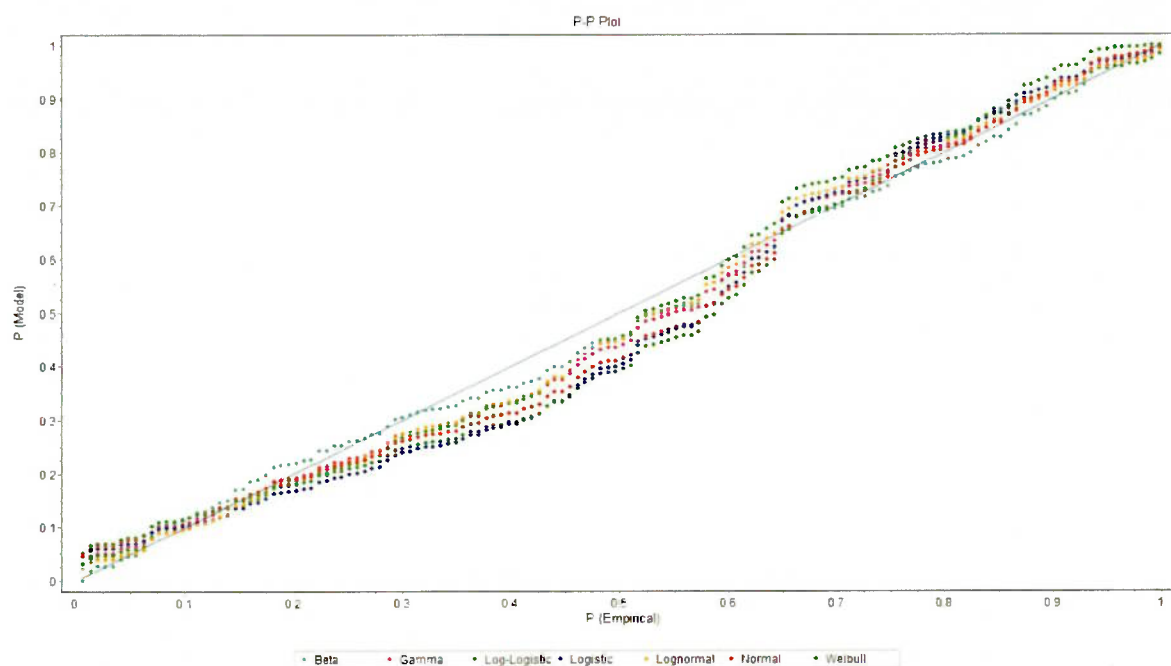


Figura 28 - PP-Plot da variável X_4 , sem outliers.

Posteriormente foram realizados os testes de hipótese e aderência para cada uma das 7 funções ajustadas à variável X_4 , considerando:

- Hipótese nula $H_0 \rightarrow$ Não rejeita H_0 quando: $X_i \in \delta$; Rejeita H_a
- Hipótese nula $H_a \rightarrow$ Não rejeita H_a quando: $X_i \notin \delta$; Rejeita H_0

O resumo dos testes de aderência para a variável X_4 é mostrado na Tabela 07. Em geral, obteve-se o melhor resultado com a distribuição beta:

Tabela 07 - Resumo dos testes de aderência variável X_4 .

Distribution	Kolmogorov Smirnov		Anderson Darling		Chi-Squared	
	Statistic	Rank	Statistic	Rank	Statistic	Rank
Beta	0,05699	1	2,3275	5	6,2028	1
Lognormal	0,07155	2	0,7833	1	6,6555	2
Log-Logistic	0,07589	3	1,1598	3	8,3028	4
Gamma	0,07663	4	0,89591	2	7,7314	3
Normal	0,09693	5	1,6935	4	11,486	5
Weibull	0,11517	6	3,4209	7	18,24	7
Logistic	0,11649	7	2,3952	6	15,204	6

A Tabela 08 apresenta o detalhamento dos resultados dos testes de hipótese para a variável X_4 utilizando a função de distribuição beta.

Tabela 08 - Detalhamento dos resultados dos testes de hipótese variável X_4 .

Kolmogorov-Smirnov		Anderson-Darling		Chi-Squared	
Sample Size	143	Sample Size	143	Deg. of freedom	7
Statistic	0,05699	Statistic	2,3275	Statistic	6,2028
Rank	1	Rank	5	Rank	1
α	5,00%	α	5,00%	α	0,05
Critical Value	0,11356	Critical Value	2,5018	Critical Value	14,067
Reject H_0 ?	No	Reject H_0 ?	No	Reject H_0 ?	No

O teste de hipótese não rejeitou a hipótese nula para nenhum dos testes de aderência com um nível de significância de 5%. Assim sendo, a distribuição beta foi a eleita para descrever a variável X_4 deste modelo com os parâmetros: $\alpha=1,4356$; $\beta=2,3756$; em um intervalo $p=\{219800 : 533010\}$. Ou seja, $X_4 \sim \text{Beta}(1,4356, 2,3756)$.

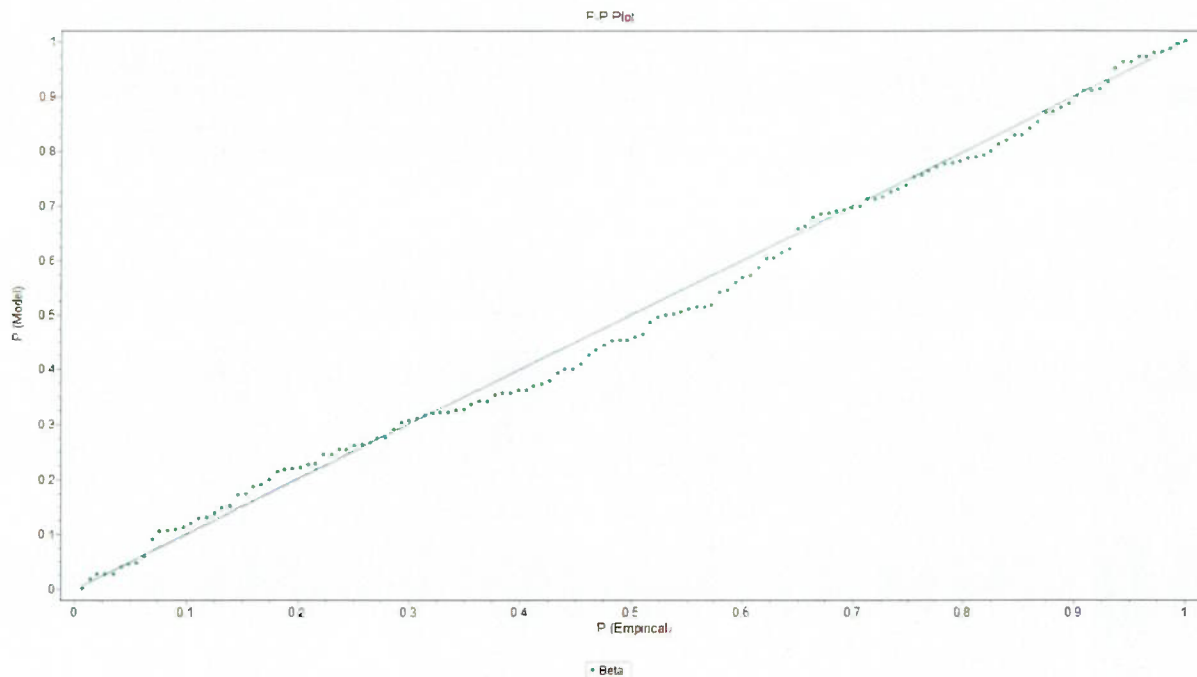


Figura 29 - PP-Plot da variável X_4 , somente com a distribuição Beta.

4.3 Simulação de Monte Carlo

Parte-se agora para a simulação de Monte Carlo com a geração de N cenários que seguem as distribuições e propriedades apresentadas anteriormente. A construção deste modelo será realizada utilizando linguagem *Python 3* com o auxílio do ambiente de desenvolvimento *Jupyter Notebook*, conforme segue:


```

# Propriedades das distribuições

# 1) X1--> Distribuição Gama
k1=28.515
teta1=7194.2
mu1=k1*teta1
sigma1=(k1*teta1**2)**0.5

# 2) X2--> Distribuição Normal
mu2=8122.1
sigma2=1285.4

# 3) X3--> Distribuição Gama
k3=42.169
teta3=1308.2
mu3=k3*teta3
sigma3=(k3*teta3**2)**0.5

# 4) X4--> Distribuição Beta
alfa4=1.4356
beta4=2.3756
mu4=alfa4/(alfa4+beta4)
sigma4=((alfa4*beta4)/(((alfa4+beta4)**2)*(alfa4+beta4+1)))**0.5

```

Figura 30 - Propriedades das distribuições das variáveis X_1 , X_2 , X_3 , X_4 .

```

N=10000 #N= número de números aleatórios
Xr=np.random.random((N,4))
X=np.zeros((N,4))
Xc=np.zeros((N,4))
for j in range(N):
    X[j,0]= (gamma.ppf(Xr[j,0],a=k1,scale=teta1)-mu1)/sigma1
    X[j,1]= (norm.ppf(Xr[j,1],mu2,sigma2)-mu2)/sigma2
    X[j,2]= (gamma.ppf(Xr[j,2],a=k3,scale=teta3)-mu3)/sigma3
    X[j,3]= (beta.ppf(Xr[j,3],alfa4,beta4)-mu4)/sigma4

X

```

```

array([[ -0.29750594, -1.02061254, -0.20083733, -1.14868479],
       [ -1.74779338, -0.12661035,  0.64215979,  0.64860288],
       [  1.22187795,  0.09412699, -0.20929249,  1.25016894],
       ...,
       [  0.337096   ,  1.98056114, -1.58961118,  0.85599229],
       [ -0.89001033,  0.1635421  ,  0.59125622, -0.78930672],
       [  1.25297935, -1.02817962,  1.25753173,  1.40927872]])

```

Figura 31 - Geração dos N cenários para as variáveis X_1 , X_2 , X_3 , X_4 .

Neste procedimento foi gerado 10.000 cenários para as variáveis X_1 , X_2 , X_3 , X_4 , conforme suas propriedades (Figura 31). Realizou-se também a padronização dos

dados por $\left(\frac{x_i - \bar{X}_i}{\sigma_i}\right)$. Entretanto, os cenários gerados possuem uma correlação totalmente diferente da matriz de correlação (Figura 11). A Figura 32 mostra a matriz de correlação gerada com os N cenários, onde percebe-se uma total desconexão entre as variáveis.

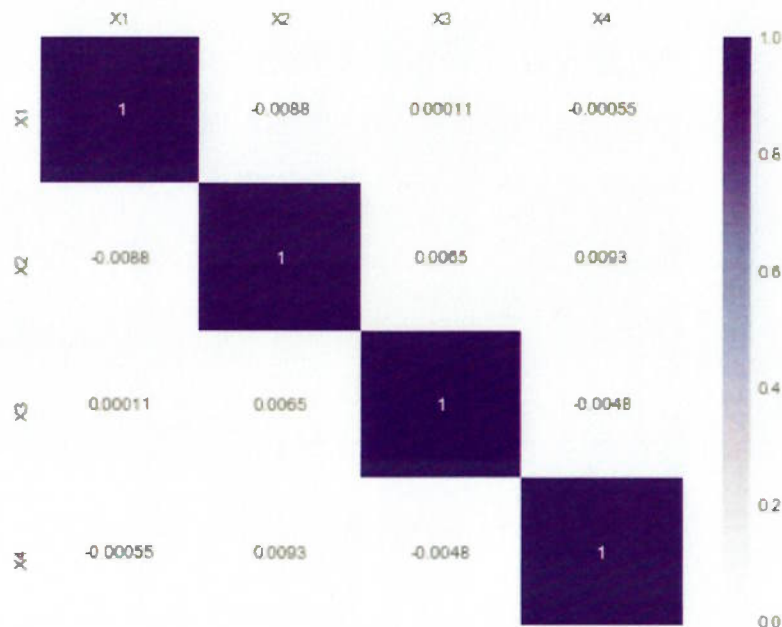


Figura 32 - Matriz de correlação gerada com os N cenários.

Assim sendo, necessita-se realizar a correlação das variáveis aleatórias conforme matriz de correlação dos dados originalmente amostrados (Figura 11). Para a realização deste procedimento será realizado a fatoração de Cholesky, conforme apresenta a Figura 33.

```
from scipy.linalg import cholesky
# Gerando a matriz de M de composição de Cholesky;
# S é a matriz de correlação original dos dados
S_chol = cholesky(S)

# produto de matrizes X x S_chol para correlacionar
# as variáveis aleatórias conforme a matriz correlação S.
Xa = X @ S_chol
```

Figura 33 - Código para correlacionar os N cenários gerados com a matriz de correlação.

Posteriormente, retorna-se à condição inicial, realizando: $(\bar{X}_i + \phi_i \cdot \sigma_i)$ e recalcula-se a matriz de correlação, conforme mostra a Figura 34.

```

Xc_=np.zeros((N,4))
for i in range(N):
    Xc_[i,0]= mu1+Xa[i,0]*sigma1
    Xc_[i,1]= mu2+Xa[i,1]*sigma2
    Xc_[i,2]=mu3+Xa[i,2]*sigma3
    Xc_[i,3]= ((IntervaloF-IntervaloI)*(mu4+Xa[i,3]*sigma4))+IntervaloI

Xc_df=pd.DataFrame(data=Xc_,columns=['X1','X2','X3','X4'])
kkk=Xc_df.corr()
sns.set_style('whitegrid')
plt.figure(figsize=(8,6))
sns.heatmap(kkk, annot= True, cmap='Purples')

```

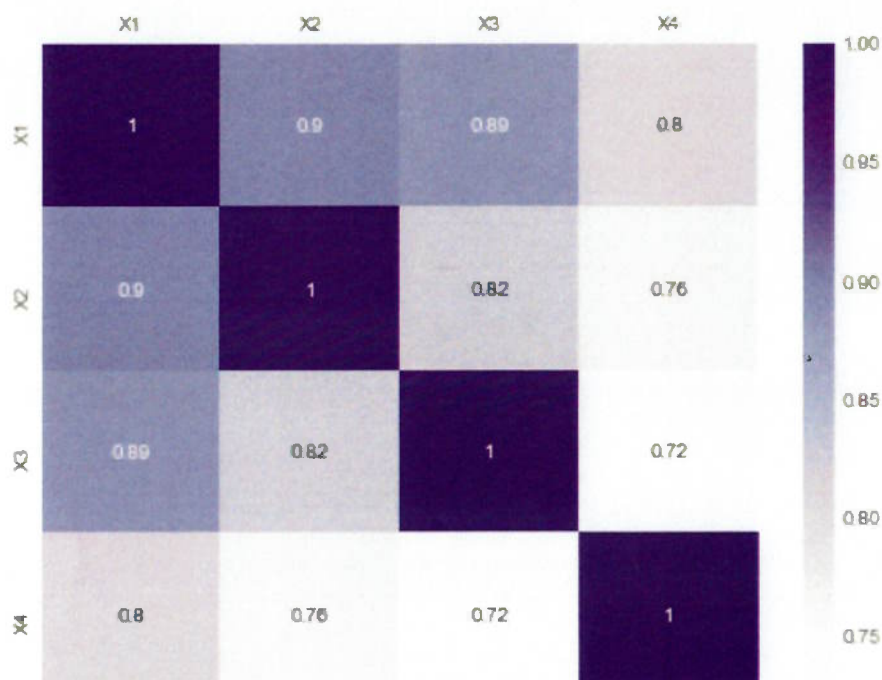


Figura 34 - Variáveis aleatórias X_i correlacionadas para N cenários.

Por fim, calcula-se o P&L (Variável X_5 : Lucro ou prejuízo) pela aplicação da equação 23, conforme apresenta a Figura 35.


```
Xc_df['P&L']=Xc_df['X4']-(Xc_df['X3']+Xc_df['X2']+Xc_df['X1'])
Xc_df.head(10)
```

	X1	X2	X3	X4	P&L
0	202716.266966	7270.795005	56659.044610	285143.722845	18497.616263
1	242964.626049	9214.261108	63680.261555	380179.802974	64320.654261
2	178053.536970	8028.190365	48430.254112	338465.750557	103953.769110
3	227564.779237	8621.768407	55291.809332	403821.721025	112343.364050
4	221495.363776	7642.829106	55422.852525	343700.926914	59139.881507
5	198933.373729	8570.957795	51212.558360	300024.304006	41307.414122
6	156153.988462	6745.611561	43761.018608	227351.150119	20690.531488
7	207457.684017	7761.613067	55666.130027	405657.159031	134771.731920
8	159553.233415	7515.043817	54768.725732	298552.219558	76715.216593
9	178960.616595	6900.311448	52409.099604	365408.757778	127138.730132
...

Figura 35 - P&L na operação de *flip* no mercado de *real estate* norte americano.

5. Apresentação e Discussão dos Resultados

Inicialmente será apresentado o histograma da variável X_5 o qual representa a distribuição de frequências dos N cenários de lucro ou prejuízo simulados para a atividade de *flip* no mercado de *real estate* norte americano (Figura 36).

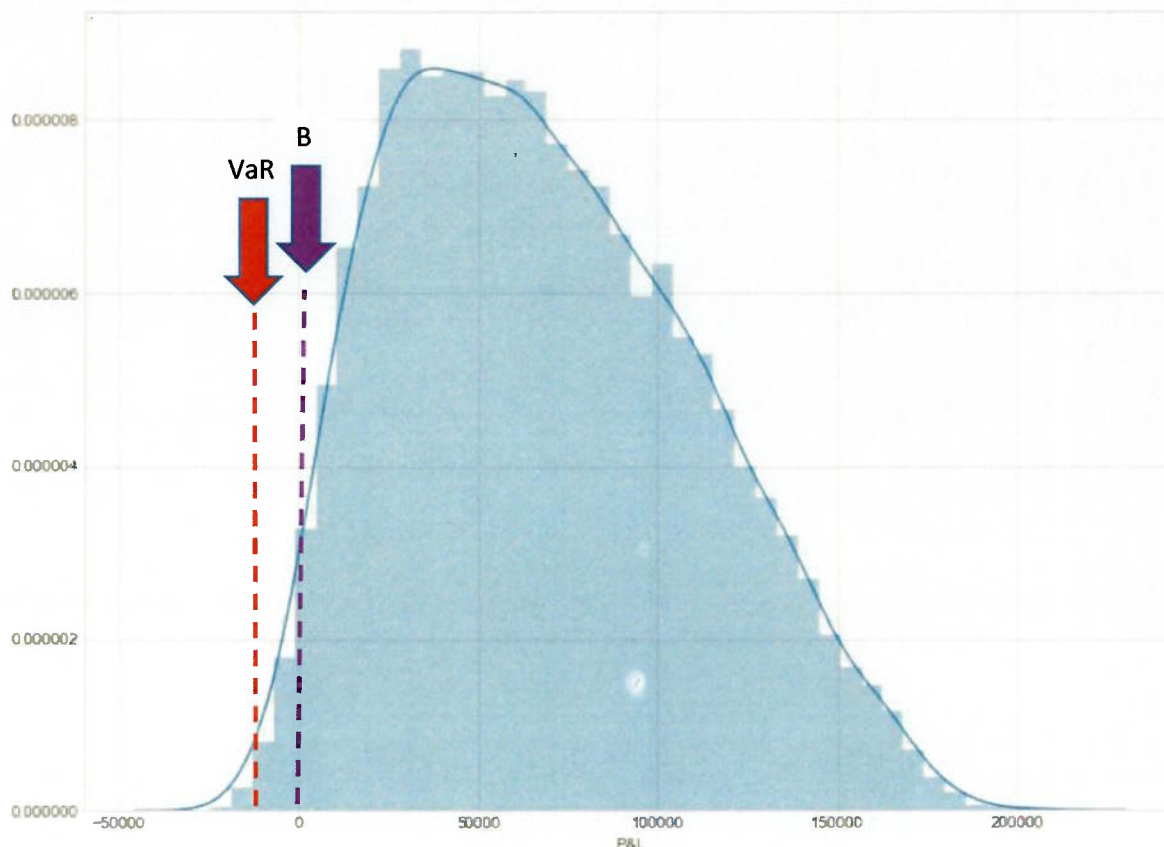


Figura 36 - Distribuição de probabilidade do P&L do investimento.

Pode-se perceber que a distribuição do P&L possui uma concentração mais acentuada em torno da média, com leve dispersão para a direita (assimetria positiva - maiores lucros). A seta roxa na Figura 36, indica o Break-Even, ou seja, o ponto de equilíbrio onde não se obtém lucro nem prejuízo. Este ponto está localizado a um percentil de 1,94% assim sendo, existe uma chance de 1,94% do investimento resultar em prejuízo ou não resultar em retorno. Consequentemente, existe a probabilidade de 98,06% do investidor realizar algum lucro na operação. O VaR a 1% obtido foi de -\$4.273,90, ou seja, o investidor neste caso estaria sujeito, com 1% de probabilidade, a aceitar um prejuízo de \$4.273,90, ou mais.

O valor esperado encontrado pelo modelo foi de \$68.241,53 com um desvio padrão de \$42.035,56. Percebe-se um alto retorno financeiro do investimento.

Vale salientar que estas análises devem ser consideradas em situações normais de mercado, o modelo não considera situações de stress (crises).

Explorando um pouco mais os resultados obtidos, agora analisaremos uma série de combinações entre as variáveis X_1 , X_2 , X_3 e X_4 com a variável X_5 , no intuito de perceber padrões que podem minimizar o risco nesta atividade de investimento. Essas combinações estão apresentadas na Tabela 09 abaixo.

Tabela 09 - Relações entre as variáveis com cenários desfavoráveis

	Considerando todos os N cenários	Considerando apenas os casos onde houve prejuízo em X_5 .	Diferença (%)
Preço médio de compra das casas (X_1)	\$205.094	\$165.512	23,9%
Gasto médio com renovação (X_3)	\$55.159	\$48.828	12,9%
$\frac{\text{Preço da Renovação } (X_3)}{\text{Preço de compra } (X_1)}$	0,276	0,287	-3,8%
$\frac{\text{Preço da venda } (X_4)}{\text{Preço de compra } (X_1)}$	1,64	1,29	27,13%

Percebe-se que os imóveis mais caros não estão relacionados com maior exposição ao risco, pelo contrário, nos cenários onde ocorreram prejuízos, em média, os preços de compra (X_1) foram 23,9% menores em comparação aos preços de compra de todos os cenários.

Outra análise que se verifica é que maiores gastos com renovação não estão diretamente ligados com uma maior exposição ao risco, nos cenários onde ocorreram prejuízos, os gastos médios com renovação foram 12,9% menores. É o que também se verifica quando se analisa a relação $\frac{X_3}{X_1}$ onde obteve-se valores praticamente iguais entre os cenários que apresentaram prejuízos com relação a todos os N cenários do modelo.

Verificou-se que na relação $\frac{X_4}{X_1}$, considerando os N cenários, o *spread* médio entre o preço de compra e o preço de venda é de 64%, enquanto para os casos em que ocorreram prejuízos esse *spread* cai para 29%.

6. Conclusões

Neste trabalho foi apresentado uma metodologia utilizando o método de Monte Carlo para a análise financeira, quantificação e gerenciamento dos riscos aplicados a um caso real de mercado. Em específico, o estudo desenvolvido analisou a atividade de *flip* no mercado de *real estate* norte americano, o qual, de forma resumida, consiste na aquisição de imóveis, renovação dos mesmos e por fim, sua venda. A amostra coletada obteve 147 unidades completas o que garantiu uma boa ideia da população. Obviamente, uma amostragem mais significativa traria ainda mais confiabilidade ao modelo, entretanto, a amostragem com 147 unidades completas foi a alcançada com os recursos disponíveis.

O modelo desenvolvido foi capaz de apresentar respostas e mostrar quantitativamente os riscos envolvidos nesta atividade de investimento, suas projeções financeiras e suas características, bem como, auxiliar o investidor na tomada de decisão. Considerando um cenário normal de mercado, em específico destaca-se que:

- a) Apesar de haver grande volatilidade com os retornos deste tipo de investimento, a atividade de flip ainda se mostra financeiramente muito interessante;
- b) Um *Break-Even* a 1,94% apresenta um risco relativamente baixo frente ao retorno esperado na atividade de \$68.241,53;
- c) Um VaR a 1% igual a -\$4.273,90, ou seja, o investidor neste caso estaria sujeito, com 1% de probabilidade, a aceitar um prejuízo de \$4.273,90, ou mais. Este risco se apresenta extremamente baixo frente ao retorno esperado na atividade de \$68.241,53;
- d) Imóveis de maior valor não significam maior exposição ao risco;
- e) Maiores gastos com renovação e reforma dos imóveis não significam maior exposição ao risco;
- f) *Spread* entre a relação preço de venda sobre preço de compra do imóvel tem grande influência no P&L deste investimento. Como pode ser observado na Figura 37, quanto mais se aproxima de uma relação $\frac{x_4}{x_1}$ igual a 1,29 maior será a sua exposição ao risco.

Assim sendo, conclui-se que o investidor deve antes de iniciar este tipo de investimento, projetar o possível valor de venda deste imóvel ou qual valor o mercado está disposto a pagar por esse tipo de imóvel renovado e buscar uma relação $\frac{X_4}{X_1}$ mais próxima da média 1.64.

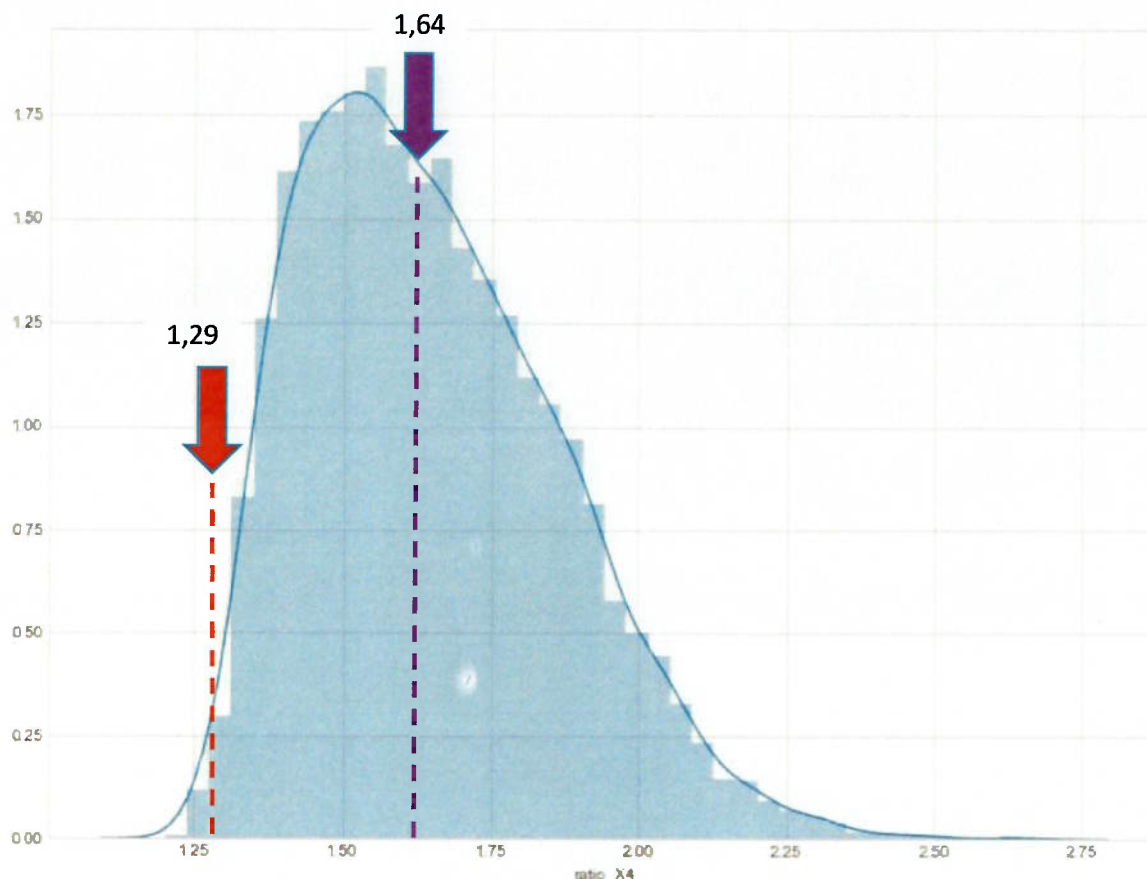


Figura 37 - Taxa $\frac{X_4}{X_1}$ nos N cenários simulados

A seta vermelha da Figura 37 indica a média da relação $\frac{X_4}{X_1}$ dos casos que obtiveram prejuízos. Já a seta roxa da Figura 37 apresenta a média da relação $\frac{X_4}{X_1}$ geral de todos os cenários gerados.

Conclui-se que, o modelo foi capaz de produzir ferramentas quantitativas que podem auxiliar o investidor na tomada de decisão, possibilitando uma melhor compreensão, controle e entendimento dos riscos envolvidos nesta atividade.

7. Trabalhos Futuros

O modelo desenvolvido neste trabalho apresentou importantes informações que auxiliarão o investidor na tomada de decisões, no controle e gerenciamento de riscos e na maximização de retornos. Entretanto, como mostrado no trabalho, o conhecimento da relação entre o preço de compra do imóvel e seu futuro preço de mercado, após renovação $\left(\frac{x_4}{x_1}\right)$ é de suma importância para a minimização de riscos e maximização de retornos. Desta forma, um modelo de avaliação imobiliária que pudesse precificar o valor futuro de venda de um imóvel renovado, considerando todas as suas características físicas (número de quartos, número de banheiros, localização, idade, etc...) seria um grande complemento a este trabalho. Auxiliando o investidor na correta escolha do imóvel, minimizando as chances de um erro humano pela falta de experiência ou conhecimento deste mercado.

Outro ponto a ser considerado e que pode ser alvo de futuros estudos é que o modelo trata apenas de uma representação estática desta atividade no mercado atual. Assim sendo, uma possível extensão deste trabalho poderia ser realizada com uma amostra mais numerosa e em um período de tempo mais amplo adicionando ao modelo a variável tempo, correlacionando o modelo com variáveis macroeconômicas.

Além disso, a mesma metodologia empregada neste trabalho pode ser diretamente aplicada para uma enorme gama de produtos de investimentos e projetos imobiliários. Muitos projetos imobiliários e projetos de infraestrutura poderiam ser melhor planejados e administrados com a utilização de ferramentas análogas as mostradas neste trabalho, minimizando imprevistos que comumente ocorrem nestas atividades. Esses imprevistos são geralmente gerados pela falta de conhecimento e tratamento das incertezas existentes.

Referências Bibliográficas

CRUDDEN, M. **A discussion of the Monte Carlo technique applied to commercial property: Examining risk in perspective.** Public Infrastructure Bulletin: Volume 1, 8.6 (2012).

FRENCH, N.; & GABRIELLI, L. **Discounted cash flow: accounting for uncertainty.** Journal of Property Investment & Finance, 23(1), 75-89, (2005).

HOESLI, M.; JANI, E.; & BENDER, A. **Monte Carlo simulations for real estate valuation.** Journal of Property Investment & Finance, 102-122, (2006).

LO, Andrew W. e MUELLER, Mark T. **WARNING: Physics Envy May Be Hazardous To Your Wealth.** Journal of Investment Management, 8.2 (2010).

MOORE, J.; WEATHERFORD, L.R. **Tomada de decisão em administração com planilhas eletrônicas.** 6.ed. Porto Alegre: Bookman Companhia Editora, 2006.

MUN, Johnathan. **Real Options Analysis: Tools and Techniques for Valuing Strategic Investments and Decisions.** 2. ed. San Francisco, California: John Wiley & Sons, Inc, 2006. 667 p.

MUN, Johnathan. **Applying Monte Carlo Risk Simulation, Strategic Real Options, Stochastic Forecasting, and Portfolio Optimization.** 2. ed. Hoboken, Nj: John Wiley & Sons, Inc, 2010. 986 p.

NAVIDI, William. **Probabilidade e Estatística para Ciências Exatas.** Porto Alegre: AMGH Editora Ltda, 2012. 604 p.

PYHRR, S. **A Computer Simulation Model to Measure Risk In Real Estate Investment.** Journal of the American Real Estate and Urban Economics Association, 48-78, (1973).

PUSTEJOVSKY, Phil. **How To Be a Real Estate Investor**. Lexington, KY: Freedom Mentor, 2017. 200 p.

REES, Michael. **Business Risk: and Simulation Modeling in Practice**. Padstow, Cornwall, Uk: Wiley & Sons, Inc, 2015. 425 p.

SUHONEN, Ville. **Using Monte Carlo Simulation to Support a Retail Real Estate Investment Decision**. 2014. 54 f. Dissertação (Mestrado) - Degree Programme In Real Estate Economics, Aalto University School Of Engineering, Greater Helsinki, Finland, 2014.

WANG, Hui. **Monte Carlo Simulation with Applications to Finance**. Providence, Rhode Island, USA: A Chapman & Hall Book, 2012. 282 p.

WILMOTT, Paul. **Paul Wilmott On Quantitative Finance**. 2. ed. Chichester, Uk: John Wiley & Sons, Inc, 2006. 1379 p.

WIKIPEDIA THE FREE ENCYCLOPEDIA (Org). **Multiple Listing Service**. 2008. Disponível em: <https://en.wikipedia.org/wiki/Multiple_listing_service#North_America> Acesso em: 26 set. 2018.