

**CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS**

**Diretoria de Pesquisa e Pós-Graduação**

**Programa de Pós-Graduação em Modelagem Matemática e Computacional**

Washington Pinheiro de Souza

Modelagem de um teste de hipóteses, baseado na convolução da distribuição Binomial e a distribuição Normal, com aplicações na indústria automotiva, medicina veterinária e no índice IPCA do Banco Central.

**Belo Horizonte**

**02 de maio de 2019**

Washington Pinheiro de Souza

Modelagem de um teste de hipóteses, baseado na convolução da distribuição Binomial e a distribuição Normal, com aplicações na indústria automotiva, medicina veterinária e no índice IPCA do Banco Central.

Dissertação apresentada ao Programa de Pós-graduação em Modelagem Matemática Computacional, no Centro Federal de Educação Tecnológica de Minas Gerais, como requisito parcial para a obtenção do título de Mestre em Modelagem Matemática e Computacional.

Orientador: Prof<sup>o</sup>. Dr<sup>o</sup> Thiago Gomes de Mattos.

**Belo Horizonte**  
**02 de maio de 2019**

S729m Souza, Washington Pinheiro de.  
Modelagem de um teste de hipóteses, baseado na convolução da distribuição Binomial e a distribuição Normal, com aplicações na indústria automotiva, medicina veterinária e no índice IPCA do Banco Central. / Washington Pinheiro de Souza. – 2019.  
65 f.

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Modelagem Matemática e Computacional.  
Orientador: Thiago Gomes de Mattos.  
Dissertação (mestrado) – Centro Federal de Educação Tecnológica de Minas Gerais.

1. Estatística – Modelos – Teses. 2. Convolução (Matemática) – Teses. 3. Distribuição binominal – Teses. 4. Distribuição Gaussiana – Teses. 5. Teste de hipótese estatística – Teses. I. Mattos, Thiago Gomes de. II. Centro Federal de Educação Tecnológica de Minas Gerais. III. Título.

CDD 519.6



SERVIÇO PÚBLICO FEDERAL  
MINISTÉRIO DA EDUCAÇÃO  
CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS  
COORDENAÇÃO DO PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL

***“MODELAGEM DE UM TESTE DE HIPÓTESES, BASEADO NA  
CONVOLUÇÃO DA DISTRIBUIÇÃO BINOMIAL E A DISTRIBUIÇÃO  
NORMAL, COM APLICAÇÕES NA INDÚSTRIA AUTOMOTIVA,  
MEDICINA VETERINÁRIA E NO ÍNDICE IPCA DO BANCO CENTRAL”***

Dissertação de Mestrado apresentada por **Washington Pinheiro de Souza**, em 30 de agosto de 2018, ao Programa de Pós-Graduação em Modelagem Matemática e Computacional do CEFET-MG, e aprovada pela banca examinadora constituída pelos professores:

Prof. Dr. Thiago Gomes de Mattos (Orientador)  
Centro Federal de Educação Tecnológica de Minas Gerais

Prof. Dr. Luiz Argel Poveda Calvião  
Centro Federal de Educação Tecnológica de Minas Gerais

Prof. Dr. Arthur Rodrigo Bosco de Magalhães  
Centro Federal de Educação Tecnológica de Minas Gerais

Prof. Dr. Fábio Rocha da Silva  
Centro Federal de Educação Tecnológica de Minas Gerais

Visto e permitida a impressão,

Prof. Dr. José Geraldo Peixoto de Faria  
Coordenador do Programa de Pós-Graduação em  
Modelagem Matemática e Computacional

## RESUMO

Em vários segmentos que necessitam de um orçamento prévio e/ou uma previsão de receitas, é comum a realização de estimativas. Tal procedimento é feito embasado em modelos estatísticos. A distribuição Binomial é um modelo estatístico discreto que determina as chances de sucesso ou fracasso da ocorrência de um determinado número de eventos em um espaço amostral limitado. A distribuição Normal é o modelo estatístico padrão que segue a lei dos grandes números, ou seja, ocorrências aleatórias que tendem ao infinito convergem para uma média e apresentam uma variância. Reunir as características das duas distribuições em uma ferramenta única é de grande valia para se realizar estimativas e, para tanto, utiliza-se a convolução. Os testes de hipóteses são ferramentas estatísticas muito utilizadas para se determinar o a abrangência de um determinado conjunto de dados numéricos a partir da rejeição ou não da hipótese nula. Para se realizar a convolução entre as duas distribuições citadas é necessário que ambas sejam discretas ou contínuas. A distribuição Binomial foi aproximada de uma função contínua, através de ferramentas matemáticas apropriadas, e posteriormente convoluída com a distribuição normal. Foi feita a previsão da quantidade de peças defeituosas que devem ser recolhidas pela Magneti Marelli em todo o Brasil. Os valores estimados contemplaram os valores reais divulgados pela empresa com uma margem de erro. No Hospital Veterinário da UFMG, foram feitas previsões da quantidade de animais que apresentaram alteração no exame de risco cirúrgico para a cirurgia de castração no ano de 2016, sendo que, caso houvesse algum valor fora da estimativa, a hipótese nula seria rejeitada. Todos os valores mensais coletados no hospital veterinário no ano de 2016 estavam dentro da previsão feita utilizando o modelo o que fez com que a hipótese nula fosse aceita. Foi estimado o Índice Nacional de Preços ao Consumidor Amplo (IPCA) dos anos de 2011 a 2016 com base nos dados de 1996 a 2010 divulgados pelo IBGE. Boa parte dos valores observados estavam dentro da faixa prevista pelo método com proximidade da média, exceto por um, fazendo com que a hipótese nula fosse rejeitada. Pela aplicação nos diversos setores citados com respectivos resultados satisfatórios, o método proposto se mostrou eficaz e multidisciplinar, ou seja, pode ser aplicado em diversas áreas.

Palavras chave: Modelo. Convolução. Distribuição. Binomial. Normal. Testes. Hipóteses.

## ABSTRACT

In several segments requiring a prior budget and / or a revenue forecast, estimates are common. This procedure is based on statistical models. The Binomial distribution is a discrete statistical model that determines the chances of success or failure of the occurrence of a certain number of events in a limited sample space. The normal distribution is the standard statistical model that follows the law of large numbers, that is, random occurrences that tend to infinity converge to a mean and present a variance. Gathering the characteristics of the two distributions in a single tool is of great value to make estimates and, for that, the convolution is used. The hypothesis tests are statistical tools widely used to determine the comprehensiveness of a given set of numerical data from the rejection or not of the null hypothesis. In order to carry out the convolution between the two distributions mentioned it is necessary that both are discrete or continuous. The Binomial distribution was approximated to a continuous function, using appropriate mathematical tools, and later convoluted with the normal distribution. The quantity of defective parts that must be collected by Magneti Marelli throughout Brazil has been forecast. The estimated values included the actual values disclosed by the company with a margin of error. At the Veterinary Hospital of UFMG, estimates were made of the number of animals that presented alterations in the surgical risk exam for castration surgery in 2016, and if there was any value outside the estimate, the null hypothesis would be rejected. All monthly values collected at the veterinary hospital in the year 2016 were within the prediction made using the model which made the null hypothesis acceptable. The National Broad Consumer Price Index (IPCA) for the years 2011 to 2016 was estimated based on data from 1996 to 2010 released by IBGE. Most of the observed values were within the range predicted by the method with proximity of the average, except for one, causing the null hypothesis to be rejected. Due to the application in the several sectors mentioned with satisfactory results, the proposed method proved effective and multidisciplinary, that is, it can be applied in several areas.

Keywords: Model. Convolution. Distribution. Binomial. Normal. Tests. Hypotheses.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Reta de eventos $X + Y = t$ em um espaço bidimensional $(x, y)$ .....	13
Figura 2 – Curva característica da distribuição Normal.....	14
Figura 3 – Região Crítica em um teste de hipóteses unicaudal a esquerda.....	18
Figura 4 – Superposição das curvas $P_1(x)$ , $P_2(x)$ e $P_3(x)$ sobre a distribuição Binomial com $p = 0,15$ e $N = 10$ .....	24
Figura 5 – Convolução entre a distribuição Normal e a aproximação contínua da distribuição Binomial.....	25
Figura 6 – Gráfico da convolução entre a distribuição normal e a aproximação contínua da distribuição binomial para $N = 10$ , $p = 0,15$ , $\mu = 20$ e $\sigma = 2$ .....	26
Figura 7 – Gráfico da convolução entre a distribuição normal e a aproximação contínua da distribuição binomial para $N = 20$ , $p = 0,27$ , $\mu = 5,4$ e $\sigma = 3,17$ .....	28
Figura 8 – Representação gráfica da distribuição mensal de animais cujo exame de risco cirúrgico para a realização da cirurgia de castração, no ano de 2016, apresentaram alteração.....	40

## LISTA DE TABELAS

Tabela 1 – Tipos de erros que podem ocorrer em um teste de hipóteses.....	17
Tabela 2 – Valores de parâmetros calculados para três curvas diferentes de $P(x)$ .....	24
Tabela 3 – Quantidade de peças defeituosas por lote de peças.....	27
Tabela 4 – Quantidade de articulações axiais defeituosas recolhidas por estado.....	30
Tabela 5 – Valores dos parâmetros a serem utilizados para realizar estimativas.....	32
Tabela 6 – Estratificação dos dados coletados no hospital veterinário da UFMG para o ano de 2015.....	36
Tabela 7 – Estimativa de exames mensais que venham a apresentar alteração em 2016 baseado nos dados de 2015.....	38
Tabela 8 – Estratificação dos dados coletados no hospital veterinário da UFMG para o ano de 2016 com a estimativa da quantidade de animais cujo exame de risco cirúrgico para castração apresentou alteração.....	39
Tabela 9 – Valores percentuais referentes ao POF atual e anterior.....	42
Tabela 10 – Valores percentuais referentes ao IPCA acumulado de 1996 à 2010.....	43
Tabela 11 – Valores dos parâmetros da aproximação contínua da distribuição Binomial para $p = 0,0652$ e $N = 100$ .....	44
Tabela 12 – Valores percentuais referentes ao IPCA acumulado de 2011 à 2016.....	45



## LISTA DE ABREVIATURAS E SIGLAS

CA	Coeficiente Angular.
CL	Coeficiente Linear.
CMM	COFAP Magneti Marelli.
EA	Exames que apresentaram alteração.
EEA	Faixa estimada de exames que apresentarão alteração
ER	Exames realizados.
FIMM	Fábrica Italiana Magneti Marelli.
IBGE	Instituto Brasileiro de Geografia e Estatística.
IPCA	Índice Nacional de Preços ao Consumidor Amplo.
QAC	Quantidade de animais aptos a realizarem a cirurgia de castração.
SpA	Società per Azioni.
VAR	Variância.

## SUMÁRIO

1. INTRODUÇÃO.....	10
1.1. Projeções e Valor Futuro.....	10
1.2. Modelo Estatístico.....	11
2. OBJETIVOS.....	19
3. METODOLOGIA.....	20
3.1. Aproximação da Distribuição Binomial de uma Distribuição Contínua.....	20
3.2. A Convolução da Distribuição Normal com a Aproximação da Distribuição Binomial Contínua.....	25
4. RESULTADOS E DISCUSSÕES.....	29
4.1. Aplicação em Número de Peças Defeituosas na Indústria Automotiva.....	29
4.2. Aplicação em Número de Animais que Estão Aptos a Realizar uma Cirurgia de Castração pela Análise do Exame de Risco Cirúrgico.....	33
4.3. Aplicação em Índice IPCA no Brasil.....	40
5. CONCLUSÃO.....	46
REFERÊNCIAS.....	48
APÊNDICE A: Cálculo de $k$ .....	50
APÊNDICE B: Obtendo as relações entre $\alpha$ e $\beta$ através do cálculo da esperança de $P(x)$ .....	53
APÊNDICE C: Obtendo as relações entre $\alpha$ e $\beta$ através do cálculo da variância de $P(x)$ .....	56
APÊNDICE D: Cálculo da Esperança e da Variância da Distribuição Normal.....	60
APÊNDICE E: Cálculo da Esperança e da Variância da Distribuição Binomial.....	62
APÊNDICE F: Cálculo da Convolução entre a Aproximação Contínua da Distribuição Binomial e a Distribuição Normal.....	64

## 1. INTRODUÇÃO

### 1.1. Projeções e Valor Futuro

Projeções são métodos estatísticos utilizados para especular valores futuros através de um histórico de ocorrências numéricas. O objetivo principal de uma metodologia para projetar não é acertar, mas ter uma ideia de um cenário para que administradores de um determinado evento possam tomar decisões com antecedência hábil (CAVALCANTE, 1998). Projeções são muito utilizadas no mercado financeiro, em bolsas de valores, na indústria e na ciência. A partir de valores projetados em análises numéricas de dados dentro de um histórico, é possível criar orçamentos, prever receitas e estabelecer metas e perspectivas necessárias para que uma empresa seja gerida de maneira eficiente e rentável, por exemplo.

Há uma grande variedade de modelos estatísticos que podem ser utilizados na previsão de resultados de eventos aleatórios. Tais modelos podem ser discretos ou contínuos. Como exemplo de modelo discreto, pode-se citar a distribuição Geométrica, que avalia a chance da primeira ocorrência de um evento bem sucedido, cuja chance de sucesso  $p$  é conhecida, a partir de uma seleção aleatória dentro de um espaço amostral de tamanho  $N$ . Como exemplo de modelo contínuo, pode-se destacar a distribuição de Weibul, que é utilizada para modelar tempos de vida. Como exemplo prático de aplicação desse tipo de modelagem estatística, é possível citar o controle de qualidade em uma montadora automotiva, onde é comum a necessidade de se estimar a quantidade de peças defeituosas dentro de um lote. Sabendo-se que um número  $J$  de lotes com  $T$  peças foram recebidos anteriormente com uma quantidade  $X$  de peças defeituosas por lote, tem-se que  $X$  é uma variável aleatória, cuja distribuição pode ser caracterizada quantitativamente a partir da sua média  $\mu$  e da sua variância  $\sigma^2$ .

Em um outro exemplo, um hospital veterinário recebe pacientes que eventualmente necessitam de uma intervenção cirúrgica. Porém, para que uma cirurgia possa ser executada, é necessário que haja uma resposta positiva de uma avaliação fisiológica, a qual é feita a partir de exames de risco cirúrgico. Estimar quantos exames apresentarão alteração, em função do histórico de ocorrência de exames com tais características, é importante para que o hospital possa se programar em termos de quadro de funcionamento, de plantões, de horários e de quantidade de materiais e medicamentos a ser mantidos em estoque. Para tanto, é possível obter uma projeção analisando

estatisticamente a variável aleatória  $Y$ , que corresponde ao número de pacientes que necessitam mas não podem se submeter a uma intervenção cirúrgica.

Em todos os exemplos, a natureza das variáveis envolvidas é estocástica, ou seja, não se sabe quantas peças defeituosas existirão em um lote de peças e não há como se determinar quantos pacientes não poderão realizar uma cirurgia devido às suas condições fisiológicas. Não obstante, há uma necessidade de se estabelecer projeções a fim de se realizar planejamentos. Dentro desse contexto, procura-se uma distribuição estatística que leve em consideração parâmetros quantificáveis, de maneira a viabilizar uma projeção, criando um modelo estatístico aceitável para tal aplicação.

## 1.2. Modelo Estatístico

Neste trabalho propomos um modelo estatístico que leva em consideração a chance de sucesso ou de insucesso de se observar um determinado número de ocorrências dentro de uma sequência de eventos com uma média e uma variância. As características desse tipo de processo podem ser descritas por uma distribuição Normal (Gaussiana) ou por uma distribuição Binomial, dependendo se o espaço amostral é contínuo ou discreto, respectivamente. A proposta consiste em um modelo que reúna aspectos dessas duas distribuições através de uma operação de convolução.

A convolução é um operador linear que, a partir de duas funções dadas, resulta numa terceira que mede a área subentendida pela superposição das mesmas em função do deslocamento existente entre elas (BRACEWELL, 2000). Ao executar a convolução entre duas funções ou sequências, o resultado carrega as características principais dos seus geradores.

Para que duas funções possam ser convoluídas é necessário que ambas sejam discretas ou contínuas. Uma sinal discreto não pode processar um sinal contínuo e vice versa. Para cada situação a convolução é tratada de uma maneira. Funções discretas são convoluídas através de um somatório

$$(f * g)(k) = h(k) = \sum_{j=0}^k f(j)g(k-j) \quad 1$$

onde  $f$  e  $g$  são sequências de tamanho  $n$  e a fórmula fornece o  $k$ -ésimo elemento do resultado.

Funções contínuas têm sua convolução dada através da integral:

$$(f * g)(t) = h(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau \quad 2$$

Em que  $f$  e  $g$  são funções contínuas de  $t$ . A convolução apresenta as seguintes propriedades:

- Comutatividade:

$$a[n] * b[n] = b[n] * a[n]$$

- Associatividade:

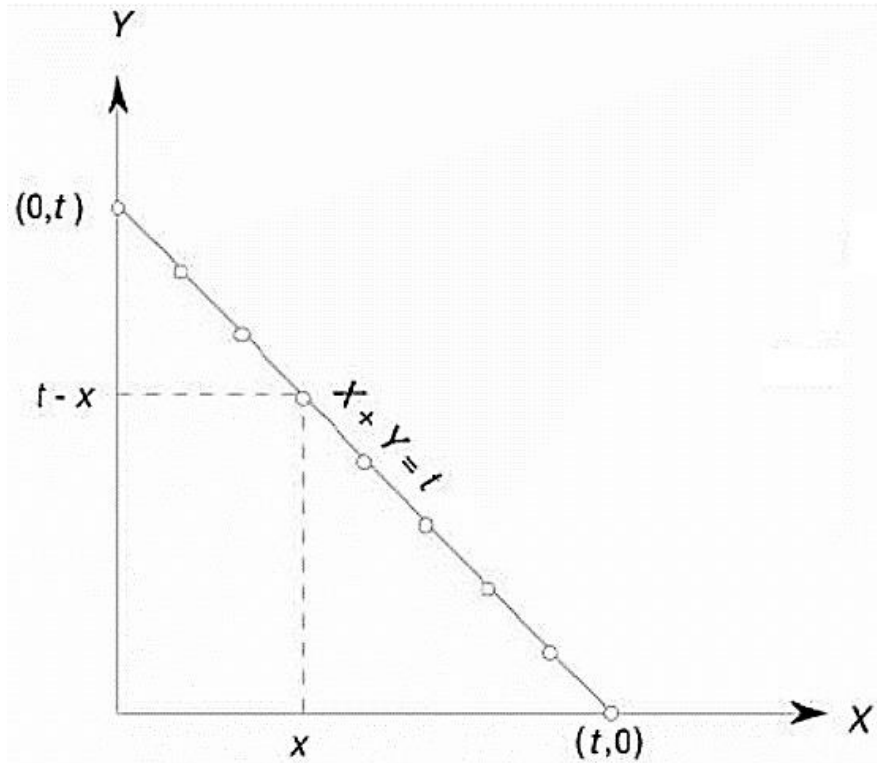
$$(a[n] * b[n]) * c[n] = a[n] * (b[n] * c[n])$$

- Distributividade:

$$a[n] * (b[n] + c[n]) = a[n] * b[n] + a[n] * c[n]$$

Na convolução discreta, a soma de todos os valores que compõem a função resultante é igual ao produto entre a soma de todos os valores dos pontos de cada função geradora. Na convolução contínua, a área abaixo do gráfico da função resultante é igual ao produto entre a área abaixo de cada gráfico das funções geradoras.

Na estatística, a convolução representa a soma de duas variáveis. Dados os eventos independentes  $X$  e  $Y$ , considerando a probabilidade do evento  $Z = X + Y = t$ , em um espaço de eventos bidimensional  $(x,y)$ , estes eventos são representados por todos os pontos na reta  $X + Y = t$ , que é demonstrada na Figura 1.



**Figura 1:** Reta de eventos  $X + Y = t$  em um espaço bidimensional  $(x, y)$ .

A probabilidade de  $Z$  pode ser obtida pela adição dos pontos nessa reta, sendo:

$$P(Z = t) = \sum_{x=0}^t P(X = x, X + Y = t) = \sum_{x=0}^t P(X = x)(Y = t - X) \quad 3$$

A relação demonstrada em 3 representa uma convolução.

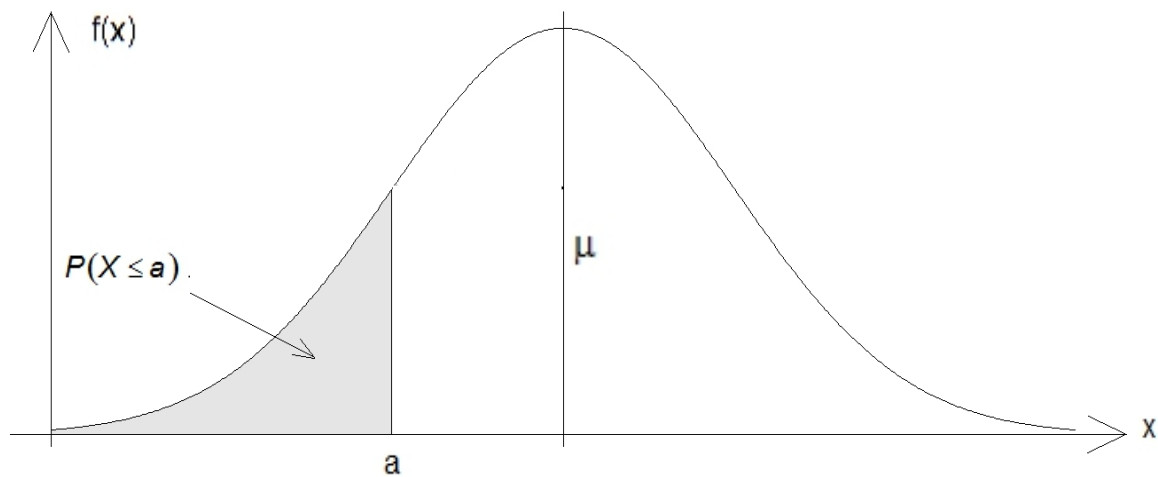
Uma variável aleatória, cujos valores são expressos em uma escala contínua, é chamada de variável aleatória contínua (CORREA, 2003). É possível construir modelos matemáticos para variáveis contínuas através da função densidade de probabilidade, que é uma função indicadora da probabilidade nos possíveis valores de  $x$ . Um modelo estatístico muito utilizado para variáveis contínuas é a distribuição Normal, ou distribuição Gaussiana.

A distribuição Normal tem grande importância em probabilidade e estatística em decorrência do Teorema do Limite Central (DANTAS, 2000), o qual afirma que, obtendo-se uma série de resultados experimentais de natureza independente, o ajuste que melhor distribuirá tais valores dentro de uma mesma densidade de probabilidades seguirá uma distribuição normal com média  $\mu$  e variância  $\sigma^2$ . Esta afirmativa decorre também da lei

dos grandes números, a qual prediz que uma amostra muito grande terá seus valores convergindo para uma média, que por sua vez estará cada vez mais próxima do valor esperado (esperança).

A densidade de probabilidade da distribuição Normal é definida no intervalo  $-\infty < x < +\infty$ .

$$f(X) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad 4$$



**Figura 2:** Curva característica da distribuição Normal.

Sendo  $P(x \leq a) = \int_{-\infty}^a \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$ .

O termo  $\sqrt{2\pi\sigma^2}$ , que se encontra no denominador da função de densidade de probabilidades da distribuição Normal, normaliza  $f(x)$ , ou seja, é o valor da área abaixo do gráfico da distribuição normal desde  $-\infty$  até  $+\infty$ .

A probabilidade de ocorrência de um evento, utilizando a distribuição Normal, se dá de maneira cumulativa, ou seja, para determinar a chance de que  $a$  ocorra deve-se calcular a área abaixo da curva que vai de  $-\infty$  até  $a$  (Figura 2). Caso  $a$  esteja em um ponto do eixo  $x$  maior que  $\mu$ , a chance da sua ocorrência é a área abaixo do gráfico que vai de  $a$  até  $+\infty$ .

Na estatística, os termos utilizados para caracterizar as distribuições de probabilidades são os momentos. Os primeiro, segundo, terceiro e quarto momentos caracterizam a esperança, dispersão (variância), assimetria e curtose, respectivamente.

A esperança matemática de uma variável aleatória é o valor médio esperado de uma experiência se ela for repetida muitas vezes (DANTAS, 2000). Por tal definição, a esperança está diretamente associada à lei dos grandes números e ao teorema do limite central. Em um experimento no qual todos os eventos possuam a mesma probabilidade, o valor esperado é a média aritmética (SILVA, 2011).

Para uma variável aleatória discreta  $x$ , cujos valores são  $x_1, x_2, x_3, \dots, x_i$ , e suas probabilidades dadas por  $p(x_i)$ , a esperança é calculada pela série:

$$E[x] = \sum_{i=1}^{+\infty} x_i p(x_i) \quad 5$$

$E[x]$  deve convergir para que a esperança exista.

Para uma variável aleatória contínua  $x$ , cujo a densidade de probabilidade é dada por  $f(x)$ , a esperança é obtida pela integral:

$$E(x) = \int_{-\infty}^{+\infty} x f(x) dx \quad 6$$

Em uma curva  $f(x)$ , de distribuição de probabilidades contínua, a esperança coincide com o ponto de máximo de  $f(x)$ . Como demonstrado no Apêndice D, a esperança da distribuição normal é a média.

Já a variância de uma variável aleatória ou processo estocástico é uma medida da sua dispersão estatística (SILVA, 2011). De modo mais grosseiro, pode-se dizer que a variância quantifica o quão distante os dados estão da esperança dentro de uma distribuição estatística.

Sendo  $var(x)$  a variância e  $E(x)$  a esperança de um modelo de distribuição estatística, logo:

$$var(x) = E((x - E(x))^2) \quad 7$$



A expressão apresentada em 7 pode ser traduzida como a esperança do desvio quadrado médio em torno da esperança  $E(x)$ . Dela decorre que:

$$var(x) = E(x^2) - E^2(x) \quad 8$$

O Apêndice D mostra que a variância da distribuição normal é:

$$var(x) = \sigma^2 \quad 9$$

A variância é o quadrado do desvio padrão ( $\sigma$ ), que por sua vez indica o quão dispersos estão os dados em torno da média. O desvio padrão pode ser amostral ou populacional. Um baixo desvio padrão indica que os dados tendem a estar próximos da esperança, caso o desvio padrão apresente um alto valor, os dados tendem a ficar muito dispersos em torno da média. Ao contrário da variância, o desvio padrão pode ser representado graficamente na curva de densidade de probabilidades. Os pontos  $\mu - \sigma$  e  $\mu + \sigma$  coincidem com os pontos de inflexão do gráfico da distribuição normal.

Uma variável  $x$  definida sobre um espaço amostral discreto é dita uma variável aleatória discreta (CORREA, 2003).

Uma variável aleatória tem distribuição Binomial quando a ela está associado um carácter booleano tal que a ocorrência de sucesso em um ensaio se dá com probabilidade  $p$ , enquanto a ocorrência de fracasso se dá com probabilidade  $q = 1 - p$ . A probabilidade de que  $x$  assuma um determinado valor  $x = a$  a partir da repetição de  $N$  ensaios independentes é:

$$P(X = a) = \binom{N}{a} (1 - p)^{N-a} (p)^a \quad 10$$

A esperança e a variância da distribuição Binomial, que são calculadas no Apêndice E, é dada por:

$$E(x) = np \quad 11$$

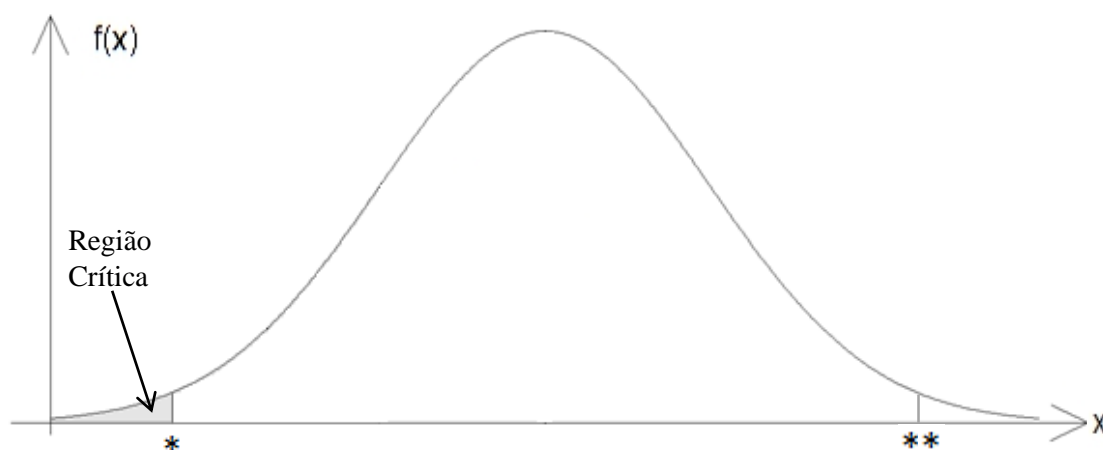
$$var(x) = np(1 - p) \quad 12$$

O teste de hipóteses é um procedimento estatístico no qual uma decisão é tomada a partir da aceitação ou rejeição de uma hipótese nula ( $H_0$ ) (FRANKLIN et al, 1998). A hipótese nula ( $H_0$ ) é a hipótese assumida como verdadeira para construção de um teste e se trata de uma teoria, efeito ou alternativa que será testada. Quando a hipótese nula não possui evidências estatísticas, fato este assumido após a realização do teste, a hipótese alternativa ( $H_a$ ) é considerada. Mesmo quando um teste é realizado, e decide-se rejeitar ou não  $H_0$ , existe a chance de que a decisão tomada não seja a correta. Para esses casos, são definidas as ocorrências de dois tipos de erros: o erro tipo I, que ocorre quando rejeita-se a hipótese nula e ela é verdadeira, e o erro tipo II, que ocorre quando rejeita-se a hipótese alternativa e ela é verdadeira. Para sintetizar o que foi descrito, o quadro abaixo é apresentado:

	<b><math>H_0</math> é verdadeira</b>	<b><math>H_0</math> é falsa</b>
<b>Hipótese nula <math>H_0</math> é rejeitada</b>	Erro tipo I	Não há erro
<b>Hipótese nula <math>H_0</math> não é rejeitada</b>	Não há erro	Erro tipo II

**Tabela 1:** Tipos de erros que podem ocorrer em um teste de hipóteses.

Em um teste de hipóteses existe um nível de significância, que é a maior probabilidade que se permite cometer o erro tipo I. Tal nível é denotado por  $\alpha$  e o seu valor é sempre pré-estabelecido antes da realização do teste. Com o nível de significância é possível determinar qual a região crítica do teste, que consiste no conjunto de valores, assumidos pela variável aleatória, para o qual a hipótese nula é rejeitada (BUSSAB, 2004). Para a determinação da região crítica, além do nível de significância, é necessário se determinar qual tipo de teste será realizado ou qual distribuição de probabilidades mais se adequa ao problema em análise.



**Figura 3:** Região Crítica em um teste de hipóteses unicaudal a esquerda.

Se o teste for unicaudal à esquerda (Figura 3), a região crítica compreenderá  $-\infty < x \leq *$ , com  $\int_{-\infty}^* f(x)dx = \alpha$ , se o teste for unicaudal à direita, a região crítica compreenderá  $** \leq x < +\infty$ , com  $\int_{**}^{+\infty} f(x)dx = \alpha$  e se o teste for bicaudal, a região crítica compreenderá  $-\infty < x \leq * \cup ** \leq x < +\infty$ , com  $\int_{-\infty}^* f(x)dx + \int_{**}^{+\infty} f(x)dx = \alpha$ . Se uma amostra, cujo teste foi feito utilizando uma amostragem da mesma natureza, apresentar valores dentro da região crítica a hipótese nula é rejeitada.

Além do uso da região crítica, um teste de hipóteses pode ser realizado utilizando a probabilidade de significância ou o  $p$ -valor. O  $p$ -valor é definido como a probabilidade de se obter uma estatística de teste igual ou mais extensa que a estatística observada a partir de uma amostra (BUSSAB, 2004).

O procedimento básico para a realização de um teste de hipóteses é:

- Escolher  $H_0$  e  $H_a$  com base no problema proposto;
- Determinar a melhor distribuição estatística para testar a hipótese nula a partir das informações disponíveis no problema;
- Determinar o nível de significância (os mais comuns são 5% e 1%);
- Se o teste for pela região crítica, deve-se determinar a região crítica, retirar uma amostra da população, fazer os cálculos para determinar o valor da estimativa de teste e caso tal valor pertencer à região crítica, a hipótese nula deverá ser rejeitada;
- Se o teste for pelo  $p$ -valor, deve-se calcular o  $p$ -valor e caso tal valor for menor que o nível de significância, a hipótese nula deverá ser rejeitada;

- Quando a hipótese nula não puder ser rejeitada, o procedimento pode ser repetido com diferentes valores para o nível de significância, visando uma maior precisão à decisão.

A convolução entre duas distribuições de probabilidade irá criar um espaço que contém a soma das duas distribuições inicialmente escolhidas, cujos parâmetros de ambas contemplarão o novo espaço obtido. Para a convolução entre a distribuição Binomial e a distribuição Normal, o espaço criado assumirá valores de probabilidades para variáveis aleatórias independentes que possuem uma média, uma variância, uma chance de sucesso e um espaço amostral limitado. Dentro desse espaço, pode-se determinar uma região para o qual uma hipótese nula é aceita com uma maior exatidão, pois com o aumento no número de parâmetros em uma análise faz com que a mesma se torne mais abrangente e mais assertiva. Para o presente trabalho, foram escolhidas a distribuição Binomial e a distribuição Normal, pois a primeira está relacionada à chance de sucesso ou fracasso de eventos aleatórios, o que está diretamente associada às aplicações a serem apresentadas neste texto, e a segunda é uma distribuição padrão para descrever uma tendência central com um grau de dispersão, outro fator muito importante para as análises que serão feitas.

## **2. OBJETIVOS**

O presente trabalho tem por objetivo desenvolver uma ferramenta para realização de um teste de hipóteses, bem como realizar estimativas que contemplem, dentro de uma região de aceitação de uma hipótese nula, situações de natureza aleatória que apresentem uma média, uma variância e uma chance de sucesso dentro de um espaço amostral limitado. Tal ferramenta terá como base a convolução da distribuição Binomial com a distribuição normal. Para tanto serão seguidos os seguintes passos:

- 1) Aproximar a distribuição Binomial por uma distribuição contínua;
- 2) Realizar a convolução entre essa aproximação e a distribuição normal;
- 3) Ajustar dados reais, a partir do modelo proposto, a fim de se obter uma relação capaz de estimar valores futuros. De posse do modelo obtido a partir da convolução entre a distribuição Normal e a aproximação contínua da distribuição Binomial, são aplicados os parâmetros para cada situação distinta, obtém-se um intervalo que contém 95% do espaço de distribuição de probabilidades conjunto das duas distribuições que sofreram a convolução e aplica-se tal intervalo em

torno da média para se estimar uma região crítica que possa ser aplicada para um teste de hipóteses. Serão considerados dados oriundos das seguintes situações:

- a) Validar o método aplicando à distribuição do número de animais cujo exame de risco cirúrgico impossibilitou (por apresentar alterações) a realização da cirurgia de castração em um determinado ano a partir da comparação das estimativas obtidas com valores conhecidos;
  - b) Aplicar o método ao número de peças defeituosas em um conjunto de lotes de peças automotivas comercializadas, visando estimar valores para a quantidade de peças defeituosas que ainda devem ser recolhidas. Comparar as projeções com valores já conhecidos;
  - c) Estimar a inflação anual acumulada no Brasil através do índice IPCA. Comparar estimativa com valores já conhecidos;
- 4) Analisar os resultados e verificar se o método é válido.

### 3. METODOLOGIA

Inicialmente, será feita a aproximação da distribuição Binomial de uma função contínua e após a conclusão de tais procedimentos, será realizada a convolução entre esta e a distribuição Normal. Com a função resultante da convolução, será estabelecido um espaço que contemple 95% da soma das distribuições para as variáveis aleatórias em estudo e, a partir desta, serão determinados os limites do espaço citado, sendo que estes comporão uma região crítica em torno da média, na distribuição Normal, no qual a hipótese nula será aceita.

Tentou-se aproximar a distribuição Normal de uma distribuição discreta, para que pudesse esse feita a convolução entre tal função e a distribuição Binomial, porém, além de ter que limitar a distribuição Binomial, o que fere o seu caráter específico de distribuição de probabilidades, observou-se uma alteração significativa nos valores de probabilidade, muito em decorrência de uma alteração na dispersão da distribuição proposta. Em decorrência deste fato, neste trabalho serão feitas apenas operações que envolvem funções contínuas.

#### 3.1. Aproximação da distribuição Binomial de uma distribuição contínua

A distribuição Binomial será aproximada de uma função contínua, cujo domínio está compreendido entre 0 e  $N$  e a esperança e a variância deverão ser iguais aos da distribuição Binomial discreta. Para tanto, será considerada a função  $P(x)$  à seguir:

$$P(x) = \frac{e^{\alpha\sqrt{x(N-x)} + \beta x}}{k} \quad 13$$

O termo  $\sqrt{x(N-x)}$  garante que o domínio de  $P(x)$  esteja restrito ao intervalo que vai de 0 a  $N$ ,  $\alpha$  e  $\beta$  são parâmetros que assumirão valores associados à esperança e à variância de  $P(x)$  e  $k$  é o termo que normaliza a aproximação contínua da distribuição Binomial, sendo assim:

$$k = \int_0^N e^{\alpha\sqrt{x(N-x)} + \beta x} dx \quad 14$$

O cálculo de  $k$  foi feito utilizando o método de Simpson e o método das coordenadas polares. Tais cálculos são mostrados no Apêndice A. Utilizando o método de Simpson, foi obtida a seguinte expressão:

$$k \approx \frac{N\pi e^{\frac{N}{2}(\alpha + \beta)}}{3} \quad 15$$

Através do método das coordenadas polares, chegou-se na expressão a seguir:

$$k = \frac{4e^{\frac{N}{2}(\alpha + \beta)}}{\alpha\beta N} \sinh\left(\beta \frac{N}{2}\right) \left[1 - \frac{2}{\alpha N} (1 - e^{-\alpha \frac{N}{2}})\right] \quad 16$$

Sendo  $\sinh\left(\beta \frac{N}{2}\right)$  o seno hiperbólico de  $\beta \frac{N}{2}$ .

A esperança de  $P(x)$  deve ser igual a esperança da distribuição Binomial, conforme é mostrado abaixo:

$$E(x) = pN = \int_0^N x \frac{e^{\alpha\sqrt{x(N-x)} + \beta x}}{k} dx \quad 17$$

O Apêndice B apresenta como foi feita a relação, a partir do cálculo da esperança, entre  $\alpha$  e  $\beta$ , utilizando o método de Simpson e o método das coordenadas polares. Após a realização dos cálculos, utilizando o método de Simpson, chegou-se a seguinte relação:

$$\frac{e^{\alpha \frac{N}{2} (\frac{\sqrt{2}}{2} - 1)}}{2} \sinh\left(\frac{\sqrt{2}N}{4} \beta\right) \approx 2p - 1 \quad 18$$

Utilizando o método das coordenadas polares, obteve-se a relação a seguir:

$$\cotgh\left(\beta \frac{N}{2}\right) - \frac{2}{\beta N} = 2p - 1 \quad 19$$

Sendo  $\cotgh\left(\beta \frac{N}{2}\right)$  a cotangente hiperbólica de  $\beta \frac{N}{2}$ .

A variância deve ser igual à variância da distribuição Binomial, conforme é mostrado abaixo:

$$Var(x) = E(x^2) - [E(x)]^2 = pN(1 - p) \quad 20$$

$$Var(x) = \int_0^N x^2 \frac{e^{\alpha \sqrt{x(N-x)} + \beta x}}{k} dx - p^2 N^2 \quad 21$$

No Apêndice C é mostrado o cálculo para se chegar às relações de  $\alpha$  e  $\beta$  com a variância, utilizando o método de Simpson e o método das coordenadas polares. Através do método de integração de Simpson, chegou-se à seguinte relação:

$$e^{\alpha \frac{N}{2} (\frac{\sqrt{2}}{2} - 1)} \approx \frac{16p(1 - p) + N(2p - 1)^2}{\sqrt{2}N \cosh\left(\frac{\sqrt{2}N}{4} \beta\right)} \quad 22$$

Substituindo 18 em 22 obtém-se:

$$\beta \approx \frac{2}{\sqrt{2}N} \ln \left[ \frac{16(1 - p) + (2p - 1)(N + 2\sqrt{2})}{16(1 - p) + (2p - 1)[N(2p - 1) - 2\sqrt{2}]} \right] \quad 23$$

O valor de  $\alpha$  é obtido isolando tal parâmetro em 22:

$$\alpha \approx \frac{4}{N(2 - \sqrt{2})} \ln \left[ \frac{\sqrt{2}N \cosh\left(\frac{\sqrt{2}N}{4}\beta\right)}{16p(1 - p) + N(2p - 1)^2} \right] \quad 24$$

Através do método das coordenadas polares obtém-se:

$$\frac{N^2(\alpha N - 6)}{2 \left[ \alpha N - 2 \left( 1 - e^{-\frac{N}{2}} \right) \right]} + \frac{12}{\alpha^2} = pN(1 - p)(N - 1) \quad 25$$

Nas equações 19 e 25,  $\beta$  e  $\alpha$  estão implícitos em uma igualdade, ou seja, para se obter os seus respectivos valores é necessário lançar números coerentes para  $p$  e  $N$  nas igualdades e, então, verificar qual o valor do parâmetro em questão é observado. Através da calculadora Hp 50G e seu programa de solução de equação (Solve Equation), é feito o cálculo para se obter o valor de  $\beta$ . Utilizando a relação proposta pela equação 25, não se chegou a valores factíveis de  $\alpha$ , pois os mesmos não seriam capazes de fornecer uma curva aproximada da distribuição Binomial quando aplicados em  $P(x)$ .

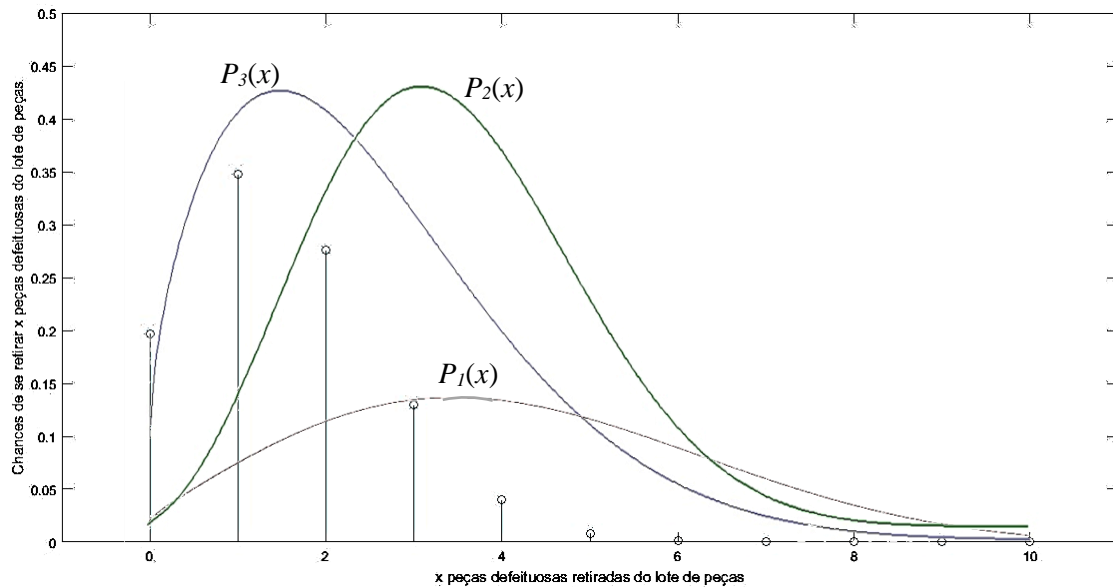
Considerando como exemplo um lote de 100 peças no qual 15 são defeituosas e se deseja saber quais as chances de que, ao se retirar 10 peças, com reposição, um número  $x$  dessas fossem defeituosas, é feita a aproximação com  $p = 0,15$  e  $N = 10$ . Como a aproximação contínua da distribuição Binomial foi feita através de dois métodos distintos de integração, será feita a comparação entre as curvas e será determinado quais os parâmetros que melhor compõem a aproximação. A tabela a seguir foi montada com os valores calculados de  $k$ ,  $\beta$  e  $\alpha$  a partir das equações 15 e 16, 19 e 23, 24 e 25, respectivamente, sendo Ref. a equação de referência para o qual o parâmetro foi calculado:



Parâmetro	$P_1(x)$		$P_2(x)$		$P_3(x)$	
	Ref.	Valor	Ref.	Valor	Ref.	Valor
$\alpha$	18	0,66	18	1,61	18	0,66
$\beta$	17	-0,21	15	-0,66	15	-0,66
$k$	12	99,35	13	515,88	13	8,8

**Tabela 2:** Valores de parâmetros calculados para três curvas diferentes de  $P(x)$ .

$P_1(x)$ ,  $P_2(x)$  e  $P_3(x)$  são as curvas que se aproximam, de maneira contínua, da distribuição Binomial a partir dos valores dos parâmetros apresentados na tabela acima. Cada curva foi traçada sobre o gráfico apresentado na Figura 4 e, então, visualmente se pôde determinar qual curva se ajusta melhor ao problema proposto.



**Figura 4:** Superposição das curvas  $P_1(x)$ ,  $P_2(x)$  e  $P_3(x)$  sobre a distribuição Binomial com  $p = 0,15$  e  $N = 10$ .

Analisando a figura acima, nota-se que a melhor aproximação, para este problema, se deu por  $P_3(x)$ . Como todo o processo se deu por métodos de integração por aproximação, e tendo em vista que em todas as etapas houve um erro de truncamento característico do método, as curvas apresentaram variações consideráveis em alguns pontos. Como  $P_3(x)$  apresentou erros, aparentemente, menores em relação à esperança e à variância, os seus parâmetros serão utilizados no restante do desenvolvimento deste trabalho.

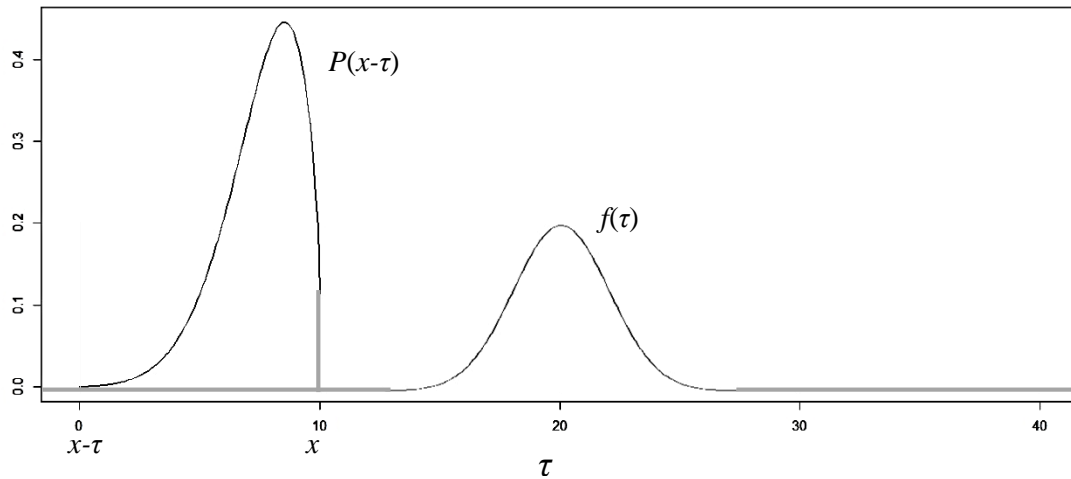
### 3.2. A convolução da distribuição normal com a aproximação da distribuição Binomial contínua

Inicialmente, considera-se a distribuição Normal e a aproximação contínua da distribuição Binomial:

$$f(X) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad 26$$

$$P(x) = \frac{e^{\alpha\sqrt{x(N-x)} + \beta x}}{k} \quad 27$$

Como a distribuição Normal possui um intervalo que vai de  $-\infty$  até  $+\infty$ , a convolução entre ambas se dará em um único estágio, conforme é demonstrado na Figura 5:

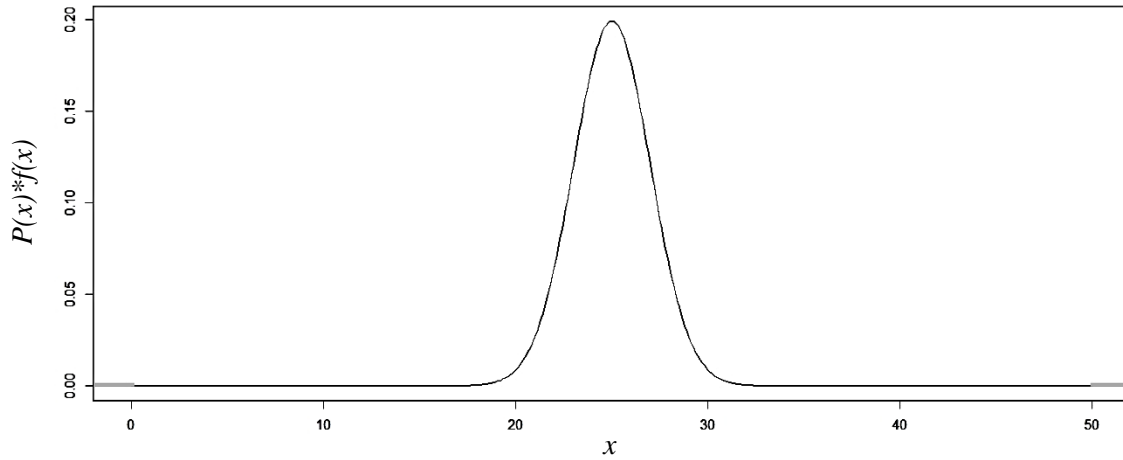


**Figura 5:** Convolução entre a distribuição Normal e a aproximação contínua da distribuição Binomial.

O Apêndice F apresenta todos os cálculos referentes à convolução entre  $f(x)$  e  $P(x)$ . Deste obtém-se a relação a seguir:

$$f(x) * P(x) \approx \sqrt{\frac{\pi}{2\sigma^2}} \frac{N e^{-\frac{(\alpha+\beta)N\sigma^2 + \left(x - \frac{N}{2}\right)^2 + (\mu-x)^2 - x^2}}{2\sigma^2}}{3k} \quad 28$$

A expressão acima pode ser traduzida, com os devidos parâmetros, no gráfico apresentado na Figura 6.



**Figura 6:** Gráfico da convolução entre a distribuição Normal e a aproximação contínua da distribuição Binomial para  $N = 10$ ,  $p = 0,15$ ,  $\mu = 20$  e  $\sigma = 2$ .

A expressão demonstrada na equação 28 deve ser normalizada. Devido à grande quantidade de parâmetros que a mesma possui, cada caso em que venha a haver a sua aplicação será normalizado individualmente.

De posse da equação 28, deve-se estabelecer um valor de  $\gamma$  no qual o espaço  $\int_{pN+\mu-\gamma}^{pN+\mu+\gamma} f(x) * P(x) dx$  compreenda 95% de todas as probabilidades de ocorrência dos eventos independentes, criando assim uma região crítica e possibilitando a realização de um teste de hipóteses.

Como exemplo, considera-se que o setor de qualidade de uma empresa recebe 10 lotes de peças contendo 20 peças cada, em que foram identificadas as seguintes quantidades de peças defeituosas por lote:

Número do lote	Quantidade de peças defeituosas
1	8
2	2
3	5
4	11
5	6
6	4
7	1
8	9
9	3
10	5
<b>Total</b>	<b>54</b>

**Tabela 3:** Quantidade de peças defeituosas por lote de peças.

Um analista deseja estimar quantas peças defeituosas poderão vir no próximo lote. Para tanto ele extrai a média e o desvio padrão dos valores apresentados na Tabela 3 e obtém  $\mu = 5,4$  e  $\sigma = 3,17$ . O parâmetro  $p$  é a razão entre o número total de peças defeituosas e o número total de peças, ou seja:

$$p = \frac{54}{10 \times 20} = \frac{54}{200} = 0,27.$$

O parâmetro  $N$  é o número de peças por lote, ou seja, 20. Com os valores obtidos é possível se chegar aos outros parâmetros da aproximação contínua da distribuição Binomial:

$$\beta = -0,1600$$

$$\alpha = 0,7021$$

$$k = 820,21$$

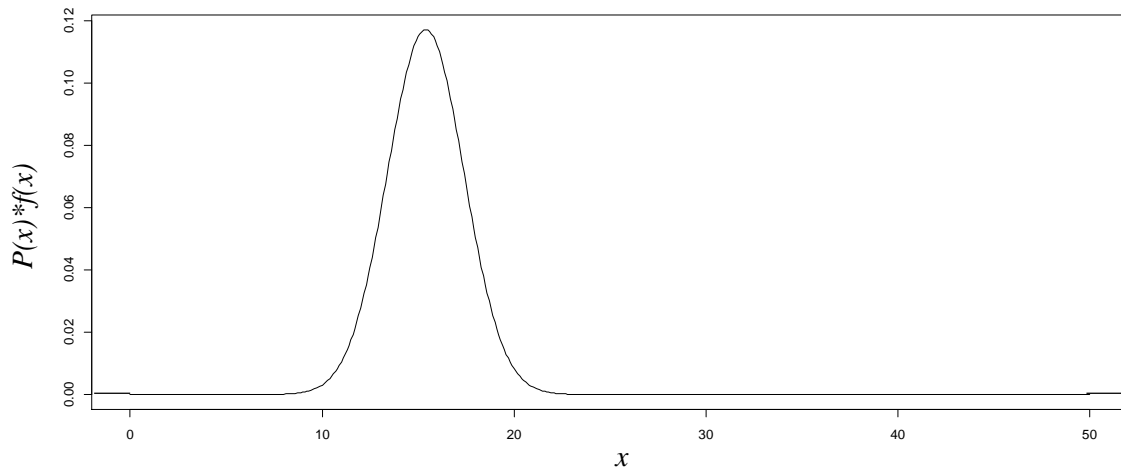
Assim, cria-se a região crítica no qual:

$$0,95c = \int_{10,8-\gamma}^{10,8+\gamma} \sqrt{\frac{\pi}{2(3,17)^2}} \frac{20e^{-\frac{(0,7021-0,16)20(3,17)^2+(x-10)^2+(5,4-x)^2-x^2}{2(3,17)^2}}}{3(820,21)} dx \quad 29$$

Sendo  $c$  uma constante de normalização, ou seja:

$$c = \int_{-\infty}^{+\infty} \sqrt{\frac{\pi}{2(3,17)^2}} \frac{20e^{-\frac{(0,7021-0,16)20(3,17)^2+(x-10)^2+(5,4-x)^2-x^2}{2(3,17)^2}}}{3(820,21)} dx \quad 30$$

$$c = 0,02436$$



**Figura 7:** Gráfico da convolução entre a distribuição Normal e a aproximação contínua da distribuição Binomial para  $N = 20$ ,  $p = 0,27$ ,  $\mu = 5,4$  e  $\sigma = 3,17$ .

Utilizando a calculadora Hp 50G, na função solve equation, obteve-se, como valor de  $\gamma$  o número 9,8194. Para que tal análise seja um teste de hipóteses, é necessário estabelecer a hipótese nula e a hipótese alternativa, logo:

$H_0$ : O processo de fabricação das peças não sofreu alteração

$H_a$ : O processo de fabricação das peças sofreu alteração

A região no qual a hipótese nula é aceita compreende  $\mu - \gamma$  até  $\mu + \gamma$ . Isso se deve ao fato de que  $\gamma$  é a “extensão”, dentro do espaço da convolução entre as distribuições, que contém 95% da soma das ocorrências entre os eventos independentes com uma média  $\mu$ , com espaço amostral  $N$  e chances de ocorrência  $p$ . A condição  $\mu - \gamma \leq N \leq \mu + \gamma$  deve ser obedecida, pois um limite de aceitação da hipótese nula deve ser coerente com objeto em estudo e nunca deve ser maior que o tamanho do espaço amostral ou menor do que

zero. Assim sendo, considerando  $y$  como a quantidade de peças defeituosas que constará no próximo lote, a hipótese nula é aceita para o intervalo:

$$5,4 - 9,8194 \leq y \leq 5,4 + 9,8194 \approx 0 \leq y \leq 15$$

Por se tratarem de quantidades de peças de valor inteiro, o resultado obtido é aproximado do número inteiro positivo mais próximo do resultado encontrado. Se houverem de 0 a 15 peças defeituosas no próximo lote, não houve alteração no processo de fabricação das peças. Caso contrário, houve mudança no processo de fabricação das peças.

## 4. RESULTADOS E DISCUSSÕES

### 4.1. Aplicação em Número de Peças Defeituosas na Indústria Automotiva

A Magneti Marelli Holding S.p.A. é uma subsidiária do grupo Fiat que produz sistemas para uso em automóveis. Foi fundada em 1919 como uma joint-venture entre a Fiat e a Ercole Marelli, sendo batizada de F.I.M.M. (Fabbrica Italiana Magneti Marelli).

A empresa atua nas mais variadas peças automotivas como sistemas elétricos, mecânicos, sistemas de exaustão, sistemas de câmbio, sistemas de suspensão entre outras tecnologias, também desenvolve sistemas em várias categorias de automobilismo.

Dentre as peças automotivas que são fabricadas pela Magneti Marelli está a articulação axial, que tem a função de ligar a cremalheira aos terminais de direção, possibilitando movimentos angulares ascendentes e descendentes na suspensão. É um item de segurança do carro e caso sua integridade seja comprometida a segurança dos ocupantes do veículo também se compromete.

Em 12 de fevereiro de 2014, iniciou-se por parte da Magneti Marelli COFAP Fabricadora de Peças LTDA (divisão Aftermarket) uma campanha de *recall* de articulações axiais fabricadas entre primeiro de agosto e 31 de outubro de 2013 na planta de Fortaleza no Ceará. A campanha consistia em recolher as peças que já haviam sido comercializadas, pois as mesmas podiam apresentar defeitos. Para tanto, os donos de veículos que possuísem tal peça em seu automóvel deveriam procurar uma oficina autorizada da fabricante do seu carro e efetuar a troca sem custos adicionais. Também foram recolhidas nos revendedores, peças pertencentes aos lotes de fabricação defeituosos. No relatório CMM (COFAP Magneti Marelli) de 08 de agosto de 2014, foi descrito que 10000 peças foram recolhidas no estoque de distribuidores e lojistas. A

quantidade de peças que foram recolhidas do consumidor final, nos estados brasileiros no qual a mesma foi vendida, até a data de 08 de agosto de 2014 é mostrada na Tabela 5.

<b>Estado</b>	<b>Peças defeituosas recolhidas</b>
São Paulo	316
Rio de Janeiro	675
Pará	28
Minas Gerais	317
Distrito Federal	7
Mato Grosso	82
Mato Grosso do Sul	48
Ceará	881
Goiás	87
Amazonas	62
Pernambuco	103
Rio Grande do Norte	190
Piauí	879
Maranhão	29
Bahia	189
Rondônia	3
Paraná	4
Rio Grande do Sul	128
Espírito Santo	12
<b>Total</b>	<b>4040</b>

**Tabela 4:** Quantidade de articulações axiais defeituosas recolhidas por estado. FONTE: CMM.

A quantidade de peças recolhidas, conforme a Tabela 4, é uma relação parcial, ou seja, até 08 de agosto de 2014 ainda restavam peças a serem recolhidas.

A proposta de estimativa que será feita a seguir, não se trata de um problema clássico de testes de hipóteses, mas se dará através do uso da ferramenta estatística desenvolvida neste texto. Será estimada a quantidade máxima de peças que ainda devem

ser trocadas, utilizando os parâmetros da convolução da distribuição normal e a aproximação contínua da distribuição binomial.

Inicialmente, obtém-se os valores de média e desvio padrão estadual, com os dados presentes na Tabela 4:

$$\mu = 212,63$$

$$\sigma = 285,46$$

O valor de  $p$  é dado por:

$$p = \frac{Q_i}{4040}. \quad 31$$

$Q_i$  é a quantidade de peças defeituosas recolhidas por estado. O valor de  $N$  é individual para cada estado e foi calculado tomando o  $N$  da relação  $\int_{pN+\mu-\gamma}^{pN+\mu+\gamma} f(x) * P(x) dx = 0,95$  como uma incógnita e  $\gamma$  como a diferença entre o valor já encontrado de peças e  $\mu + pN$ , ou seja, considerando o valor de peças já encontrado como o limite inferior no qual a hipótese nula é aceita.

Utilizando a igualdade:

$$0,95 = \int_{Q_i}^{2\mu+2pN-Q_i} \frac{\sqrt{\frac{\pi}{2\sigma^2}} N e^{-\frac{(\alpha+\beta)N\sigma^2 + \left(x-\frac{N}{2}\right)^2 + (\mu-x)^2 - x^2}}{3kc}} dx \quad 32$$

sendo

$$c = \int_{-\infty}^{+\infty} \frac{\sqrt{\frac{\pi}{2\sigma^2}} N e^{-\frac{(\alpha+\beta)N\sigma^2 + \left(x-\frac{N}{2}\right)^2 + (\mu-x)^2 - x^2}}{3k}} dx \quad 33$$

Para os valores de  $\alpha$ ,  $\beta$  e  $k$  foram consideradas as equações 24, 23 e 16, respectivamente. Os cálculos foram feitos utilizando a calculadora Hp 50G na função Solve Equation. Por se tratarem de quantidades de peças de valor inteiro, o resultado obtido é aproximado do número inteiro positivo mais próximo do resultado encontrado. Diante dessa relação, a Tabela 6 é montada e apresentada a seguir:



<b>Estado</b>	<b><math>Q_i</math></b>	<b><math>p</math></b>	<b><math>N</math></b>
São Paulo	316	0,0782	3683
Rio de Janeiro	675	0,1671	4703
Pará	28	0,0069	100
Minas Gerais	317	0,0785	3690
Distrito Federal	7	0,0017	23
Mato Grosso	82	0,0203	431
Mato Grosso do Sul	48	0,0119	198
Ceará	881	0,2181	4712
Goiás	87	0,0215	471
Amazonas	62	0,0153	282
Pernambuco	103	0,0255	626
Rio Grande do Norte	190	0,0470	1934
Piauí	879	0,2175	4716
Maranhão	29	0,0072	105
Bahia	189	0,0468	1916
Rondônia	3	0,0007	9
Paraná	4	0,0010	12
Rio Grande do Sul	128	0,0317	926
Espírito Santo	12	0,0030	38
<b>Total</b>	4040	1	28573

**Tabela 5:** Valores dos parâmetros a serem utilizados para realizar estimativas.

No documento divulgado no dia 08 de agosto de 2014 pela CMM, constava que a quantidade de produtos sujeitos a defeitos atingiu o número de 32646 peças comercializadas. Contudo, 10000 peças foram recolhidas no estoque de distribuidores e lojistas, o que se conclui que o número de peças atingidas pelo *recall* importou em 22646. O valor estimado pelo método aqui apresentado foi de 28573 peças, ou seja, o valor real que foi apresentado pelo documento está dentro do intervalo obtido e demonstrado na Tabela 5.

#### **4.2. Aplicação em Número de Animais que Estão Aptos a Realizar Cirurgia de Castração pela Análise do Exame de Risco Cirúrgico**

Dentro da rotina de um hospital veterinário estão as intervenções cirúrgicas como um método de tratamento eficaz. Normalmente, um animal que passará por uma cirurgia necessita da mesma devido à sua saúde não estar conforme deveria e medicamentos não mais surtirem efeito sobre sua patologia. Outra categoria de animais que passarão pelo referido procedimento são os que serão castrados. A castração é uma medida de profilaxia, ou seja, que visa a prevenção de doenças além de evitar a reprodução do paciente, fato esse que em um país como o Brasil chega a ser um problema de saúde pública, uma vez que existe um grande número de animais domésticos não castrados abandonados no país que se tornam uma fonte de geração e transmissão de parasitas e doenças que são nocivas aos seres humanos, como a raiva, por exemplo. Esses mesmos animais se reproduzirão, potencializando cada vez mais os riscos à saúde humana.

Para realização de qualquer procedimento cirúrgico, um animal deve passar por um controle a fim de verificar se o mesmo reúne condições fisiológicas de se submeter a cirurgia bem como de se recuperar de maneira rápida e sem agravos à sua saúde. Tal controle envolve, em procedimentos não complexos, exames de sangue, urina e um eletrocardiograma. Neste trabalho serão utilizados dados referentes somente à castração de cães e gatos, pois entende-se, a princípio, que animais que serão castrados estão saudáveis fazendo com que não sejam esperadas alterações no exame de risco cirúrgico, o que não aconteceria com um animal doente.

No exame de sangue para animais que serão castrados são avaliados o volume globular percentual, a taxa de hemoglobina, a taxa de leucócitos, a taxa de ureia e a taxa de creatinina no sangue. Cada um desses parâmetros possui uma faixa de valores aceitável e caso os valores encontrados no exame esteja fora dessa faixa o animal é considerado inapto a realizar o procedimento, dando início a um tratamento visando regularizar os referidos valores para que o paciente possa ser operado.

O volume globular é a porcentagem de hemácias no sangue e é obtido no eritrograma I, que é um dos exames realizado com a amostra do tecido hematopoiético do animal. Hemácias são unidades morfológicas da série vermelha do sangue, também designadas por eritrócitos ou glóbulos vermelhos e a sua função é transportar o oxigênio (principalmente) e o gás carbônico (CO<sub>2</sub>) (em menor quantidade) aos tecidos (VERRASTRO, 2005). As hemácias vivem por aproximadamente 120 dias. A anemia e a hidratação excessiva do animal produzem valores baixos de volume globular sendo que

a segunda pode estimular a primeira. Já a desidratação produz valores altos. O volume globular aceitável para uma cirurgia de castração deve estar entre 37 e 55% para cães e 24 e 45% para gatos. Volumes acima desta faixa podem levar o animal a uma alta frequência cardíaca durante o procedimento devido a aplicação de anestésicos e um volume abaixo desta faixa podem comprometer a vida do animal pois o mesmo não poderá perder muito sangue no decorrer da cirurgia. Gatos possuem naturalmente um volume globular menor que os cães pois a frequência cardíaca do referido felino é maior que a do cão o que resulta em uma demanda menor de glóbulos vermelhos para atender a sua necessidade de oxigênio em comparação ao canídeo (LOPES et al, 2007).

A hemoglobina é uma proteína conjugada formada de 96% de proteínas e por um grupo prostético de coloração vermelha chamado heme (4%), o qual é formado por ferro e grupamento porfínicos (LOPES et al, 2007). A hemoglobina atua de maneira conjugada com as hemácias o que permite o transporte de oxigênio pelo sistema circulatório. A taxa de hemoglobina é obtida através do exame de eritograma II. A deficiência de ingestão de ferro e anormalidades renais tornam a formação de hemoglobina deficiente por parte do organismo do animal. A faixa de hemoglobina aceitável para um cão é de 12 a 18g/dl e para um gato é de 8 a 15g/dl. As possíveis consequências de se realizar o procedimento em animais cujos valores de hemoglobina obtidos estão fora das faixas citadas são os mesmos para o caso do volume globular, bem como o porquê de gatos possuírem uma faixa de valor menor que a dos cães.

Os leucócitos ou glóbulos brancos são células produzidas pela medula óssea e atuam no sistema imunológico do animal, tendo como função a defesa do organismo do mesmo (LOPES et al, 2007). A quantidade de leucócitos por volume sanguíneo é obtida através do exame de leucograma I. A deficiência de leucócitos no sangue se deve a doenças como a leucemia (FELV nos gatos). Em cães a quantidade de leucócitos totais deve estar entre 6000 e 17000/mm<sup>3</sup> e em gatos a faixa aceitável está entre 5500 e 19500/mm<sup>3</sup>. Abaixo dos referidos valores o animal ficará com a recuperação comprometida devido a susceptibilidade de infecções e acima dos mesmos valores o animal pode apresentar uma disfunção hormonal que pode alterar atuações de anestésicos e medicamentos administrados no paciente. Geralmente cães e gatos possuem valores de leucócitos por volume sanguíneo muito próximos.

A ureia é um composto orgânico, de fórmula geral (NH<sub>2</sub>)<sub>2</sub>CO que se forma principalmente no fígado, sendo filtrada pelos rins e eliminada na urina ou pelo suor, onde é encontrada abundantemente. Constitui o principal produto terminal do metabolismo

proteico nos mamíferos. Em quantidades menores, está presente no sangue, na linfa, nos fluidos serosos proveniente da decomposição das células do corpo e também das proteínas dos alimentos. Uma das principais provas bioquímicas de função renal é a ureia, cujos valores obtidos no exame de urinálise são utilizados para detectar se há doença renal ou não no animal bem como analisar a síntese hepática do metabolismo do paciente. Quando existem lesões hepáticas extensas é possível observar a diminuição do nível de ureia sérica, desde que não haja lesões renais retardando sua eliminação (LOPES et al, 2007). A faixa de ureia admitida para cães é de 20 a 56mg/dl e em gatos é de 40 a 60mg/dl. A atividade renal em gatos é maior que em cães, por isso a produção de ureia por dl de sangue nos felinos é maior.

A creatinina é um produto da degradação da fosfocreatina (creatina fosforilada) no músculo, e é geralmente produzida em uma taxa praticamente constante pelo corpo. Através da medida da creatinina do sangue, do volume urinário das 24 horas e da creatinina urinária é possível calcular a taxa de filtração glomerular, que é um parâmetro utilizado em exames médicos para avaliar a função renal. Assim como a ureia a creatinina é uma prova bioquímica de função renal e é obtida também através do exame de urinálise. A faixa admitida para a creatinina no sangue de cães é de 0,5 a 1,5mg/dl e de gatos é de 0,8 a 1,8mg/dl. As alterações apresentadas no exame de urinálise (que envolve tanto a ureia quanto a creatinina) podem representar lesões nos rins e no fígado, o que por sua vez caracteriza uma dificuldade de síntese e excreção de compostos químicos como medicamentos e anestésicos que serão utilizados durante e depois da cirurgia.

Foram coletados junto ao laboratório de patologia clínica do hospital veterinário e da escola de veterinária da UFMG 962 exames de risco cirúrgico de castração para cães e gatos no ano de 2015 e mais 632 para o ano de 2016. Os dados foram separados por ocorrência mensal, tanto do número de exames quanto os exames que apresentaram alterações que impediram a realização do procedimento cirúrgico conforme a tabelas a seguir:

<b>Estratificação dos exames realizados em 2015</b>			
<b>Mês</b>	<b>Animal</b>	<b>Exames Realizados</b>	<b>Exames que apresentaram alteração</b>
<b>Janeiro</b>	Cão	65	34
	Gato	4	3
<b>Fevereiro</b>	Cão	51	28
	Gato	3	2
<b>Março</b>	Cão	75	42
	Gato	13	9
<b>Abril</b>	Cão	83	34
	Gato	18	16
<b>Maiο</b>	Cão	58	26
	Gato	19	16
<b>Junho</b>	Cão	88	40
	Gato	16	8
<b>Julho</b>	Cão	42	23
	Gato	12	8
<b>Agosto</b>	Cão	40	19
	Gato	18	11
<b>Setembro</b>	Cão	75	36
	Gato	26	18
<b>Outubro</b>	Cão	95	38
	Gato	24	17
<b>Novembro</b>	Cão	56	23
	Gato	30	17
<b>Dezembro</b>	Cão	33	18
	Gato	18	11
<b>Total</b>	<b>#</b>	<b>962</b>	<b>497</b>

**Tabela 6:** Estratificação dos dados coletados no hospital veterinário da UFMG para o ano de 2015.

Será feita uma análise dos exames que apresentaram alteração em função do número total de exames realizados (cães e gatos somados), para se estipular a quantidade de animais

que apresentarão alteração no exame de risco cirúrgico para a castração, a partir dos dados obtidos no ano de 2015. A verificação dos valores futuros se dará com os dados obtidos no ano de 2016. Inicialmente, consideram-se as hipóteses:

$H_0$ : A demanda de exames está normal

$H_a$ : A demanda de exames está anormal

O foco é determinar uma faixa de exames que podem apresentar alteração no ano de 2016, sem que a hipótese nula seja rejeitada, ou seja, dentro de uma condição considerada normal. Uma condição considerada normal é aquela em que não há um surto de alguma doença, relacionada a cães e gatos, ou uma variação significativa da população dos referidos animais ou até mesmo variações econômicas dos proprietários que possibilite custear um tratamento no hospital veterinário da UFMG dentro de um período. Primeiro, será utilizada a convolução entre a aproximação contínua da distribuição Binomial e a distribuição Normal para se estimar um valor de  $N$  (assim como foi feito na subseção 4.1), ou seja, uma quantidade esperada de animais que venham a realizar o exame de risco cirúrgico. De posse do valor de  $N$ , será criada a região de aceitação da hipótese nula.

Os valores de média e desvio padrão mensal dos casos de exames de risco cirúrgico alterados são calculados e mostrados abaixo:

$$\mu = 41,42$$

$$\sigma = 9,99$$

O valor de  $p$  para a análise a ser feita é independente para cada mês, pois a quantidade de exames realizados em um mês não interfere na quantidade de exames a ser realizado nos próximos meses. Tal parâmetro consiste na razão entre a quantidade mensal de exames que apresentaram alteração e a quantidade total de exames realizados em um mês:

$$p = \frac{Q_i}{Q_E} \tag{34}$$

$Q_i$  é a quantidade de exames que apresentaram alteração em um dado mês,  $Q_E$  é a quantidade de exames que foram realizados no referido mês. A estimativa consistirá em encontrar o valor de  $N$  que representa a quantidade máxima de exames no qual é possível

se obter o número mensal de exames com alteração que foram apresentados na Tabela 6 sob as condições de média, desvio padrão e chances mensais de se obter um exame alterado dentre o total de exames realizados. Sendo  $\gamma = \mu + pN - Q_i$ , a faixa no qual a hipótese nula será aceita é  $\mu - \gamma \leq Q_i \leq \mu + \gamma$ . Considerando as igualdades abaixo:

$$0,95 = \int_{Q_i}^{2\mu+2pN-Q_i} \sqrt{\frac{\pi}{2\sigma^2}} \frac{N e^{-\frac{(\alpha+\beta)N\sigma^2 + \left(x-\frac{N}{2}\right)^2 + (\mu-x)^2 - x^2}}{3kc}} dx \quad 35$$

$$e \ c = \int_{-\infty}^{+\infty} \sqrt{\frac{\pi}{2\sigma^2}} \frac{N e^{-\frac{(\alpha+\beta)N\sigma^2 + \left(x-\frac{N}{2}\right)^2 + (\mu-x)^2 - x^2}}{3k}} d \quad 36$$

Para os valores de  $\alpha$ ,  $\beta$  e  $k$  foram utilizadas as equações 24, 23 e 16, respectivamente. Os cálculos foram feitos através da calculadora Hp 50G na função Solve Equation. Os números obtidos foram arredondados para o número inteiro positivo mais próximo. Após os cálculos serem feitos, a Tabela 7 é montada e apresentada a seguir com os resultados:

Mês	$Q_E$	$Q_i$	$p$	$N$	$\mu - \gamma$	$\mu + \gamma$
Janeiro	69	37	0,5362	58	6	77
Fevereiro	54	30	0,5556	45	5	78
Março	88	51	0,5795	79	5	78
Abril	101	50	0,4950	83	9	74
Maio	77	42	0,5454	65	7	76
Junho	104	48	0,4615	80	11	72
Julho	54	31	0,5741	46	5	78
Agosto	58	30	0,5172	46	6	77
Setembro	101	54	0,5346	87	7	75
Outubro	119	55	0,4622	94	12	71
Novembro	86	40	0,4651	66	9	74
Dezembro	51	29	0,5686	44	4	79
<b>Total</b>	962	497	-	793	86	909

**Tabela 7:** Estimativa de exames mensais que venham a apresentar alteração em 2016 baseado nos dados de 2015.

Como o tamanho de  $N$  é menor do que  $\mu + \gamma$  e considerando  $N$  como a quantidade de exames que serão realizados em cada mês, o limite superior da região que aceita a hipótese nula será igual a  $N$ . Tendo sido feitas as estimativas, a Tabela 8 é apresentada comparando os valores estimados com os valores ocorridos no ano de 2016. Para uma melhor leitura da Tabela 8, considerar:

**EER:** Estimativa da quantidade de exames a serem realizados;

**EEA:** Faixa estimada de exames que apresentarão alteração;

**ER:** Exames realizados;

**EA:** Exames que apresentaram alteração;

**H<sub>0</sub> aceito:** A hipótese nula foi aceita?

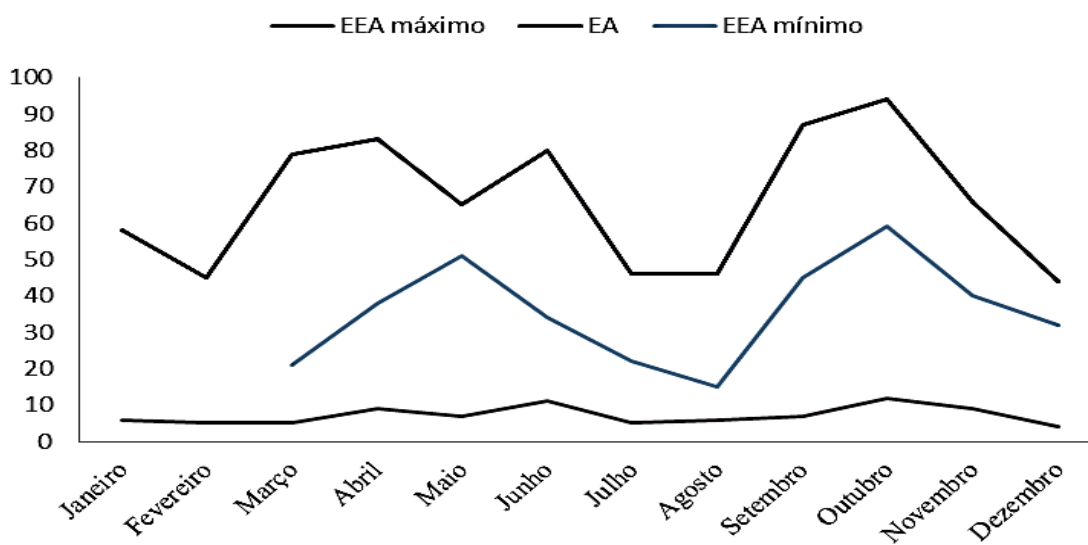
<b>Estratificação dos exames realizados em 2016</b>					
<b>Mês</b>	<b>EER</b>	<b>EEA</b>	<b>ER</b>	<b>EA</b>	<b>H<sub>0</sub> aceito</b>
Janeiro	58	6-58	-	-	-
Fevereiro	45	5-45	-	-	-
Março	79	5-79	47	21	Sim
Abril	83	9-83	64	38	Sim
Maio	65	7-65	93	51	Sim
Junho	80	11-80	55	34	Sim
Julho	46	5-46	35	22	Sim
Agosto	46	6-46	29	15	Sim
Setembro	87	7-87	92	45	Sim
Outubro	94	12-94	109	59	Sim
Novembro	66	9-66	58	40	Sim
Dezembro	44	4-44	45	32	Sim
<b>Total</b>	<b>793</b>	<b>86-793</b>	<b>632</b>	<b>361</b>	<b>Sim</b>

**Tabela 8:** Estratificação dos dados coletados no hospital veterinário da UFMG para o ano de 2016 com a estimativa da quantidade de animais cujo exame de risco cirúrgico para castração apresentou alteração.



Valores referentes aos exames realizados e aos exames que apresentaram alteração nos meses de janeiro e fevereiro não foram apresentados na Tabela 8 porque não foi possível obter tais dados relacionados aos referidos meses.

O gráfico da Figura 8 mostra que todos os valores referentes a quantidade de animais, cujo exame de risco cirúrgico para a realização da cirurgia de castração, no ano de 2016, apresentaram alteração, está dentro da estimativa realizada:



**Figura 8:** Representação gráfica da distribuição mensal de animais cujo exame de risco cirúrgico para a realização da cirurgia de castração, no ano de 2016, apresentaram alteração.

A estimativa pode ser utilizada para realizar orçamentos anuais de compra de materiais necessários para realização de cirurgias de castração, considerando um cenário em que a hipótese nula não é rejeitada, ou seja, em caráter de normalidade das operações dentro do hospital veterinário.

#### 4.3. Aplicação em Índice IPCA Anual no Brasil

As aplicações até aqui utilizadas envolveram dados de natureza discreta. Os valores a serem trabalhados na aplicação a seguir são percentuais, ou seja, não inteiros.

O Índice Nacional de Preços ao Consumidor Amplo (INPC), também conhecido como IPCA, é o indicador oficial do Governo Federal para aferição das metas inflacionárias. O IPCA foi desenvolvido pelo Instituto Brasileiro de Geografia e Estatística (IBGE) em 1979 e começou a ser divulgado a partir de Janeiro de 1980. O IBGE também produz alguns indicadores de variação de preços com objetivos específicos, como é o caso atualmente do IPCA-E (Índice Nacional de Preços ao

Consumidor Amplo Especial) e do IPCA-15 (Índice Nacional de Preços ao Consumidor Amplo 15). Assim como o IPCA tradicional, ambos verificam as variações dos gastos das pessoas que ganham de um a quarenta salários mínimos nas regiões metropolitanas de Belém, Belo Horizonte, Curitiba, Fortaleza, Porto Alegre, Recife, Rio de Janeiro, Salvador, São Paulo, Goiânia e Distrito Federal.

O IPCA foi instituído inicialmente com a finalidade de corrigir as demonstrações financeiras das companhias de capital aberto. Desde junho de 1999, é o índice utilizado pelo Banco Central do Brasil para o acompanhamento dos objetivos estabelecidos no sistema de metas de inflação, sendo considerado o índice oficial de inflação do país.

O Índice IPCA é oficialmente utilizado pelo Governo Federal para verificar se os preços praticados no Brasil estão de acordo com as metas de inflação estabelecidas pelo Banco Central do Brasil.

A inflação é um processo de elevação de preços que ocorre sempre que há procura maior do que a capacidade de uma economia produzir determinado bem ou serviço. Em resumo, a inflação pode ser de oferta, quando há escassez de produto, ou de demanda, quando a procura é maior do que a quantidade ofertada. O controle da inflação é vital para a manutenção do valor da moeda. Quando o IPCA divulgado pelo IBGE aponta que a inflação real encontra-se mais elevada do que a meta estipulada pelo Governo, o Banco Central começa a lançar mão de seus instrumentos de controle da elevação dos preços, dando início a um ciclo de alta da taxa básica de juros, a chamada taxa SELIC.

Para o cálculo do IPCA, O Sistema Nacional de Preços ao Consumidor (SNIPC) efetua a produção contínua e sistemática de índices de preços ao consumidor, tendo como unidade de coleta estabelecimentos comerciais e de prestação de serviços, concessionária de serviços públicos e domicílios (para levantamento de aluguel e condomínio). O período de coleta estende-se, em geral, do dia 01 ao dia 30 do mês de referência. A população-objetivo abrange as famílias com rendimentos mensais compreendidos entre 1 e 40 salários-mínimos, qualquer que seja a fonte de rendimentos, e residentes nas áreas urbanas das regiões. Os dados são coletados junto as famílias nas regiões metropolitanas de Belém, Fortaleza, Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, Curitiba e Porto Alegre, Brasília e Goiânia. O índice é divulgado mensalmente pelo IBGE.

A partir de Janeiro de 2012 o índice passou a ser calculado com base nos valores de despesa obtidos na Pesquisa de Orçamentos Familiares (POF) de 2008/2009. A POF é realizada a cada cinco anos pelo IBGE em todo o território brasileiro o que permite

atualizar os pesos (participação relativa do valor da despesa de um item consumido em relação à despesa total) dos produtos e serviços nos orçamentos das famílias. De julho de 2006 à dezembro de 2011 a base dos índices de preços ao consumidor era a POF de 2002/2003. A Tabela 9 apresenta os últimos valores de POF apresentados pelo IBGE:

<b>Pesos dos Grupos de Produtos e Serviços</b>		
<b>Tipos de Gasto</b>	<b>Antes de 31/12/2011</b>	<b>Depois de 01/01/2012</b>
Alimentação e bebida	23,46%	23,12%
Transporte	18,69%	20,54%
Habitação	13,25%	14,62%
Saúde e cuidados pessoais	10,76%	11,09%
Despesas pessoais	10,54%	9,94%
Vestuário	6,94%	6,67%
Comunicação	5,25%	4,96%
Artigos de residência	3,90%	4,69%
Educação	7,21%	4,37%
<b>Total</b>	<b>100%</b>	<b>100%</b>

**Tabela 9:** Valores percentuais referentes ao POF atual e anterior. Fonte: IBGE.

Diversas modalidades de investimento são indexadas pelo acúmulo do índice ao longo de 1 ano. O IPCA acumulado anual é calculado pela soma dos valores mensais ao longo dos 12 meses anteriores ao mês de aniversário do investimento.

Será feito um teste de hipóteses para o IPCA acumulado dos anos de 2011 à 2016, utilizando dados dos anos de 1996 à 2010. Para tanto, o método a ser utilizado deverá ser a convolução da distribuição normal com a aproximação contínua da distribuição Binomial. Considerando as seguintes hipóteses:

$H_0$ : A economia brasileira não sofreu grandes variações

$H_a$ : A economia brasileira sofreu grandes variações

O termo grandes variações para a economia brasileira se refere às condições mercadológicas de grande impacto, como preços de ações de grandes empresas, balança comercial, poder de compra e venda da população e etc.

A Tabela 10 apresenta os dados referentes ao IPCA acumulado dos anos de 1996 à 2010:

<b>Ano</b>	<b>IPCA acumulado (%)</b>
1996	9,56
1997	5,22
1998	1,66
1999	8,94
2000	5,97
2001	7,67
2002	12,53
2003	9,30
2004	7,60
2005	5,69
2006	3,14
2007	4,45
2008	5,90
2009	4,31
2010	5,90
<b>Total</b>	<b>97,84</b>

**Tabela 10:** Valores percentuais referentes ao IPCA acumulado de 1996 à 2010.

Fonte: IBGE.

Da Tabela 10 extrai-se os seguintes parâmetros:

$$\mu = 6,52$$

$$\sigma = 2,79$$

$$N = 100$$

O valor de  $N$  é 100, pois a característica do IPCA é percentual, ou seja, o valor máximo que este pode assumir é 100% (que ocorre quando os preços dos produtos listados na Tabela 10 dobram seu valor de um ano para o outro). O valor de  $p$  é dado por:

$$p = \frac{\mu}{N} = 0,0652.$$

Com o valor de  $p$  calculado, os demais parâmetros da aproximação contínua da distribuição Binomial são obtidos e mostrados na Tabela 11:

Parâmetro	Valor
$N$	100
$p$	0,0652
$\alpha$	0,1238
$\beta$	-0,1534
$k$	4092180,82

**Tabela 11:** Valores dos parâmetros da aproximação contínua da distribuição Binomial para  $p = 0,0652$  e  $N = 100$ .

Os valores de  $k$ ,  $\beta$  e  $\alpha$  foram obtidos utilizando as equações 16, 23 e 24, respectivamente.

$$0,95 = \int_{13,04-\gamma}^{13,04+\gamma} \sqrt{\frac{\pi}{2(2,79)^2}} \frac{100e^{-\frac{(0,1238-0,1534)100(2,79)^2+(x-50)^2+(6,52-x)^2-x^2}{2(2,79)^2}}}{3c(4092180,82)} dx \quad 38$$

$$c = \int_{-\infty}^{+\infty} \sqrt{\frac{\pi}{2(2,79)^2}} \frac{100e^{-\frac{(0,1238-0,1534)100(2,79)^2+(x-50)^2+(6,52-x)^2-x^2}{2(2,79)^2}}}{3(4092180,82)} dx = 1,73 \times 10^{14} \quad 39$$

Utilizando a função Solve Equation, da calculadora gráfica Hp 50G, e as relações

$$\int_{pN+\mu-\gamma}^{pN+\mu+\gamma} f(x) * P(x) dx = 0,95, \text{ obtém-se o valor de } \gamma:$$

$$\gamma = 4,12$$

Com o valor de  $\gamma$ , se estabelece o intervalo  $y$  para o qual a hipótese nula é aceita, sendo tal intervalo composto por:

$$2,4 \leq y \leq 10,64$$

A Tabela 12 apresenta os valores de IPCA de 2011 à 2016 e se a hipótese nula foi ou não rejeitada para cada período anual:

<b>Ano</b>	<b>IPCA acumulado (%)</b>	<b>H<sub>0</sub> aceita?</b>
2011	6,50	Sim
2012	5,83	Sim
2013	5,91	Sim
2014	6,41	Sim
2015	10,67	Não
2016	6,29	Sim

**Tabela 12:** Valores percentuais referentes ao IPCA acumulado de 2011 à 2016. Fonte: IBGE.

Na Tabela 12, é observado que somente um valor de IPCA está fora do intervalo, que é o de 2015, fazendo com que a hipótese nula seja rejeitada. Nesse caso, alguma alteração no cenário econômico brasileiro deve ser observada. De fato, tal período corresponde a um momento de grande instabilidade política no país e, de acordo com o IBGE, houve um grande aumento na taxa de desempregados no país, entre 2014 e 2015, com aumento da dívida pública e de inadimplentes no país. Os demais valores se distribuem de maneira mais uniforme em torno da média estimada.

## 5. CONCLUSÃO

O presente trabalho propunha a realização de estimativas de valores e a modelagem de um teste de hipóteses para dados de natureza aleatória, utilizando uma ferramenta matemática desenvolvida a partir da convolução entre a distribuição Binomial e a distribuição normal. Para viabilizar tal execução, foi feita uma aproximação da distribuição Binomial de uma função contínua. De posse dessa aproximação, a convolução foi realizada e devido ao fato da metodologia se basear em aproximações, observou-se um erro de aproximação característico do processo.

Como a convolução modifica o tamanho do espaço amostral da análise, deve ser feito um ajuste proporcional que condicione os parâmetros da convolução ao espaço amostral original e com tal ajuste foi possível estimar a quantidade de peças defeituosas a serem recolhidas pela Magneti Marelli em todo o Brasil, cujo valor estimado para a quantidade de peças a serem recolhidas esteve próximo ao valor real divulgado pela empresa. Para os trabalhos na patologia clínica do Hospital Veterinário da UFMG, foi estimada a quantidade mensal de animais que apresentariam alteração no exame de risco cirúrgico para cirurgias de castração, no ano de 2016, a partir de dados do ano de 2015. Para tanto, foi necessário estimar quantos exames por mês eram esperados, de maneira similar ao que foi feito para projetar a quantidade de peças defeituosas a serem recolhidas pela Magneti Marelli, e, posteriormente, foi determinado um intervalo no qual se espera que ocorram as quantidades de exames com alteração dentro de uma perspectiva de normalidade, ou seja, criou-se um intervalo para a aceitação de uma hipótese. Tendo sido feita tal análise, verificou-se que a quantidade mensal de animais cujo exame de risco cirúrgico apresentou alteração no ano 2016 ficou dentro do intervalo proposto, fazendo com que a hipótese nula seja aceita. Por fim, foi feita a estimativa do índice IPCA acumulado de 2011 a 2016 com base nos valores de IPCA dos anos entre 1996 e 2010. Ao contrário das duas aplicações feitas anteriormente, nesta o valor de  $N$  é limitado a 100, por serem dados percentuais, o que não demandou estimar o tamanho de um espaço amostral para ocorrência dos valores projetados. Como resultado, verificou-se uma distribuição homogênea dos dados em torno da média estimada, com um ponto fora do intervalo da estimativa. De um modo geral, nas três aplicações a ferramenta apresentada neste texto satisfaz a proposta inicial, uma vez que, todos os resultados esperados, ou estiveram dentro das projeções, ou foram capazes de serem justificados por uma alteração significativa.

O método aqui apresentado, pode ser aplicado em situações em que necessite fazer certas estimativas ou para avaliação de mudança no modo em que um processo se dá, quando o número de amostras é muito pequeno e não é possível se realizar uma análise de variância. Dentro desta perspectiva, podem ser feitas novas aplicações com quantidades maiores de amostras e para tempos mais longínquos, a fim de se verificar como a ferramenta proposta se comporta.



## REFERÊNCIAS

ADVFN. Indicadores IPCA. Disponível em <<https://br.advfn.com/indicadores/ipca>>. Acesso em 25 abr. 2017.

BRACEWELL, R. The Fourier Transform and It's Applications. 3rd. Edition. New York. McGraw-Hill. 2000. Cap. 3. p. 24-54.

BROWN, R. G.; GLEASON, A. M. Advanced Mathematics: Precalculus with Discrete Mathematics and Data Analysis. Evanston. McDougal Litell. 1997. 900p.

BUSSAB, W. O.; MORETTIN, P. A. Estatística Básica. 5. Ed. Saraiva. São Paulo. 2004. 537p.

CAVALCANTE, F. Como fazer "rapidamente" projeções de demonstrações financeiras. Up-To-Date®, Ano I, n. 18, São Paulo, out. 1998, 24 p. Disponível em <<http://www.cavalcanteassociados.com.br/utd/UpToDate018.pdf>>. Acesso em: 05 nov. 2016.

CASELLA, G.; BERGER, R. L. Inferência Estatística. 2. Ed. Centage Learning. São Paulo. 2010. p. 54.

COFAP. MAGENTI MARELLI. Comunicado de Recall de Peças de Reposição CMM. Mar. 2014. Disponível em <<http://www.procon.ma.gov.br/files/2012/11/RECALL-MAGNETI-MARELLI-COFAP-FABRICADORA-DE-PE%C3%87AS-LTDA.pdf>>. acesso em 25 abr. 2017.

COGNIZANT. Replacing the annual budget with Rolling Forecast. Cognizant 20-20 Insights, may. 2011, 5 p. Disponível em <<http://www.cognizant.com/InsightsWhitepapers/Replacing-the-Annual-Budget-with-Rolling-Forecasts.pdf>>. Acesso em: 05 nov. 2015.

CORMEN, T. H.; LEISERSON, C. E.; RIVEST, R. L. Introduction to Algorithms. 3. Ed. The MIT Press. 2009.

CORREA, S. M. Probabilidade e Estatística. 2. edição. PUC Minas Virtual. Belo Horizonte. 2003. p. 80-93.

CUNHA, M. T. Cálculo III. Departamento de Matemática. ICEx. UFMG. 6p. Disponível em <<http://www.mat.ufmg.br/site/wp-content/uploads/2014/07/06MudCoordenadas.pdf>>. Acesso em: 12 out. 2017.

DANTAS, C. A. B. Probabilidade: Um curso introdutório. 2. edição. Ed. USP. São Paulo. 2000. 253p.

FRANKLIN, A. G.; HARIRARAN, K. I.; RICHARD, K. B. Applied Statistics: A First Course in Inference. Prentice Hall. 1998. 461p.

IMOA (International Molybdenum Association). Practical Guidelines for the Fabrication of Duplex Stainless Steel. 2. edição. 2009. Londres. 64 p.

LOPES, S. T. A.; BIONDO, A. W.; SANTOS, A. P. Manual de Patologia Clínica Veterinária. Universidade Federal de Santa Maria. CCR. Departamento de Clínica de Pequenos Animais. 3. edição. 2007. Santa Maria. 117 p.

MIOTTO, Rafael. Fundação Nacional de Veículos Automotores. G1, jan. 2013. Disponível em <<http://g1.globo.com/carros/noticia/2013/01/brasil-fecha-2012-com-novo-recorde-de-vendas-aponta-fenabreve.html>>. Acesso em: 05 nov. 2015.

MORETTIN, L. G. Estatística Básica. 7. Ed. Vol 1. Pearson Makron Books. São Paulo. 1999. 210p.

SPERANDIO, M. M. Calculo Numérico: Características Matemáticas e Computacionais dos Métodos Numéricos. Vol único. Pearson. São Paulo. 2003.

SILVA, M. E. Uma Nota sobre Esperança Condicional e Expectativas Racionais. Universidade de São Paulo. São Paulo. 2011. Disponível em <[http://www.econ.fea.usp.br/medsilva/material/eae0308/textos/Esperanca\\_Condicional\\_e\\_ER1.pdf](http://www.econ.fea.usp.br/medsilva/material/eae0308/textos/Esperanca_Condicional_e_ER1.pdf)>. Acesso em: 12 de out. 2017.

STEWART, J. Cálculo. 4. Ed. Vol. 2. Pioneira. Thompson Learning. São Paulo. 2002. 1150p.

VERRASTRO, T. Hematologia e Hemoterapia. Ed. Atheneu. 2005. São Paulo. 303 p.

## Apêndice A: Cálculo de $k$

$$k = \int_0^N e^{\alpha\sqrt{x(N-x)} + \beta x} dx = \int_0^N e^{\alpha\sqrt{xN-x^2} + \beta x} dx = \int_0^N e^{\alpha\sqrt{\left(x-\frac{N}{2}\right)^2 + \frac{N^2}{4}} + \beta x} dx$$

Substituição trigonométrica:

$$x - \frac{N}{2} = \frac{N}{2} \operatorname{sen} \theta$$

$$x = \frac{N}{2} (1 + \operatorname{sen} \theta)$$

$$dx = \frac{N}{2} \cos \theta d\theta$$

$$0 \leq x \leq N \text{ e } -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}.$$

$$k = \frac{N}{2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\alpha\sqrt{-\frac{N^2}{4}\operatorname{sen}^2 \theta + \frac{N^2}{4}} + \beta\frac{N}{2} + \beta\frac{N}{2}\operatorname{sen} \theta} \cos \theta d\theta$$

$$k = \frac{Ne^{\beta\frac{N}{2}}}{2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha\cos \theta + \beta\operatorname{sen} \theta)} \cos \theta d\theta$$

A resolução da integral resultante se dará por dois métodos: pelo método de Simpson e pelo método da inversão das coordenadas polares.

- Pelo método de Simpson:

$$f(\theta) = e^{\frac{N}{2}(\alpha\cos \theta + \beta\operatorname{sen} \theta)} \cos \theta$$

$$a = -\frac{\pi}{2}$$

$$b = \frac{\pi}{2}$$

$$f(a) = 0$$

$$f(b) = 0$$

$$f\left(\frac{a+b}{2}\right) = e^{\alpha \frac{N}{2}}$$

$$\int_a^b f(\theta) d\theta \approx \frac{2\pi}{3} e^{\alpha \frac{N}{2}}$$

Assim:

$$k \approx \frac{N\pi e^{\frac{N}{2}(\alpha + \beta)}}{3}$$

- Pelo método das coordenadas polares inversa:

Pelo fato da integral em questão já advir de um sistema de coordenadas cartesiano, será adotado um sistema de coordenadas alternativo:

$$k = \frac{Ne^{\beta \frac{N}{2}}}{2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos \theta + \beta \sin \theta)} \cos \theta d\theta = Ne^{\beta \frac{N}{2}} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^1 e^{\frac{N}{2}(\alpha \cos \theta + \beta \sin \theta)} \cos \theta r dr d\theta$$

$$x' = \cos \theta$$

$$y' = \sin \theta$$

$$x'^2 + y'^2 = r^2, \text{ com } r = 1.$$

$$-1 \leq y' \leq 1$$

$$0 \leq x' \leq 1$$

$$k = \frac{Ne^{\beta \frac{N}{2}}}{2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos \theta + \beta \sin \theta)} \cos \theta d\theta = Ne^{\beta \frac{N}{2}} \int_0^1 \int_{-1}^1 e^{\frac{N}{2}(\alpha x' + \beta y')} x' dy' dx'$$

$$\int_0^1 \int_{-1}^1 e^{\frac{N}{2}(\alpha x' + \beta y')} x' dy' dx' = \frac{2}{\beta N} \int_0^1 (e^{\frac{N}{2}(\alpha x' + \beta)} - e^{\frac{N}{2}(\alpha x' - \beta)}) x' dx'$$

A integral resultante é resolvida pelo método de integração por partes, o que apresenta, como resultado, a seguinte relação:

$$\begin{aligned} \frac{2}{\beta N} \int_0^1 (e^{\frac{N}{2}(\alpha x' + \beta)} - e^{\frac{N}{2}(\alpha x' - \beta)}) x' dx' &= \frac{4}{\alpha \beta N^2} \left[ \left( e^{\frac{N}{2}(\alpha x' + \beta)} - e^{\frac{N}{2}(\alpha x' - \beta)} \right) \left( x' - \frac{2}{\alpha N} \right) \right]_0^1 \\ \frac{4}{\alpha \beta N^2} \left[ \left( e^{\frac{N}{2}(\alpha x' + \beta)} - e^{\frac{N}{2}(\alpha x' - \beta)} \right) \left( x' - \frac{2}{\alpha N} \right) \right]_0^1 &= \frac{4}{\alpha \beta N^2} \left( e^{\frac{N}{2}(\alpha + \beta)} - e^{\frac{N}{2}(\alpha - \beta)} \right) \left( 1 - \frac{2}{\alpha N} (1 - e^{-\frac{\alpha N}{2}}) \right) \end{aligned}$$

$$\frac{e^{\beta \frac{N}{2}} - e^{-\beta \frac{N}{2}}}{2} = \sinh \left( \beta \frac{N}{2} \right)$$

$$\frac{4}{\alpha \beta N^2} \left( e^{\frac{N}{2}(\alpha + \beta)} - e^{\frac{N}{2}(\alpha - \beta)} \right) \left( 1 - \frac{2}{\alpha N} (1 - e^{-\frac{\alpha N}{2}}) \right) = \frac{8}{\alpha \beta N^2} e^{\frac{\alpha N}{2}} \sinh \left( \beta \frac{N}{2} \right) \left( 1 - \frac{2}{\alpha N} (1 - e^{-\frac{\alpha N}{2}}) \right)$$

$$k = \frac{N e^{\beta \frac{N}{2}}}{2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos \theta + \beta \sin \theta)} \cos \theta d\theta = \frac{N e^{\beta \frac{N}{2}}}{2} \cdot \frac{8}{\alpha \beta N^2} e^{\frac{\alpha N}{2}} \sinh \left( \beta \frac{N}{2} \right) \left( 1 - \frac{2}{\alpha N} (1 - e^{-\frac{\alpha N}{2}}) \right)$$

$$k = \frac{4 e^{\frac{N}{2}(\alpha + \beta)}}{\alpha \beta N} \sinh \left( \beta \frac{N}{2} \right) \left[ 1 - \frac{2}{\alpha N} (1 - e^{-\frac{\alpha N}{2}}) \right]$$

## Apêndice B: Obtendo as relações entre $\alpha$ e $\beta$ através do cálculo da esperança de $P(x)$

$$E(x) = pN = \int_0^N x \frac{e^{\alpha\sqrt{x(N-x)} + \beta x}}{k} dx = \frac{1}{k} \int_0^N x e^{\alpha\sqrt{\left(x-\frac{N}{2}\right)^2 + \frac{N^2}{4}} + \beta x} dx$$

Utilizando a mesma substituição trigonométrica do Apêndice A:

$$x - \frac{N}{2} = \frac{N}{2} \operatorname{sen} \theta$$

$$x = \frac{N}{2} (1 + \operatorname{sen} \theta)$$

$$dx = \frac{N}{2} \cos \theta d\theta$$

$$0 \leq x \leq N \text{ e } -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}.$$

$$\frac{1}{k} \int_0^N x e^{\alpha\sqrt{\left(x-\frac{N}{2}\right)^2 + \frac{N^2}{4}} + \beta x} dx = \frac{1}{k} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{N}{2} (1 + \operatorname{sen} \theta) e^{\frac{N}{2}(\alpha \cos \theta + \beta \operatorname{sen} \theta + \beta)} \frac{N}{2} \cos \theta d\theta$$

$$E(x) = \frac{N e^{\beta \frac{N}{2}}}{2k} \cdot \frac{N}{2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos \theta + \beta \operatorname{sen} \theta)} \cos \theta d\theta + \frac{N^2 e^{\beta \frac{N}{2}}}{4k} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos \theta + \beta \operatorname{sen} \theta)} \operatorname{sen} \theta \cos \theta d\theta$$

$$k = \frac{N e^{\beta \frac{N}{2}}}{2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos \theta + \beta \operatorname{sen} \theta)} \cos \theta d\theta$$

$$E(x) = \frac{N}{2} + \frac{N^2 e^{\beta \frac{N}{2}}}{4k} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos \theta + \beta \operatorname{sen} \theta)} \operatorname{sen} \theta \cos \theta d\theta = pN$$

$$\frac{N e^{\beta \frac{N}{2}}}{2k} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos \theta + \beta \operatorname{sen} \theta)} \operatorname{sen} \theta \cos \theta d\theta = 2p - 1$$

- Caso a integral acima seja resolvida pelo método de Simpson tradicional, o resultado será 0. Para contornar tal problema, será utilizado o método de Simpson composto com  $n = 4$  e:

$$f(\theta) = e^{\frac{N}{2}(\alpha \cos \theta + \beta \operatorname{sen} \theta)} \cos \theta \operatorname{sen} \theta$$

$$h = \frac{\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)}{4} = \frac{\pi}{4}$$

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f(\theta) d\theta \approx \frac{\pi}{12} \left[ f\left(-\frac{\pi}{2}\right) + 4f\left(-\frac{\pi}{4}\right) + 2f(0) + 4f\left(\frac{\pi}{4}\right) + f\left(\frac{\pi}{2}\right) \right]$$

$$f\left(-\frac{\pi}{2}\right) = f(0) = f\left(\frac{\pi}{2}\right) = 0$$

$$f\left(-\frac{\pi}{4}\right) = -\frac{e^{\frac{\sqrt{2}N}{4}(\alpha-\beta)}}{2}$$

$$f\left(\frac{\pi}{4}\right) = \frac{e^{\frac{\sqrt{2}N}{4}(\alpha+\beta)}}{2}$$

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f(\theta) d\theta \approx \frac{\pi}{3} e^{\frac{\sqrt{2}N}{4}\alpha} \left[ \frac{e^{\frac{\sqrt{2}N}{4}\beta} - e^{-\frac{\sqrt{2}N}{4}\beta}}{2} \right]$$

$$\frac{e^{\frac{\sqrt{2}N}{4}\beta} - e^{-\frac{\sqrt{2}N}{4}\beta}}{2} = \sinh\left(\frac{\sqrt{2}N}{4}\beta\right)$$

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f(\theta) d\theta \approx \frac{\pi}{3} e^{\frac{\sqrt{2}N}{4}\alpha} \sinh\left(\frac{\sqrt{2}N}{4}\beta\right)$$

$$\frac{Ne^{\beta\frac{N}{2}}}{2k} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha\cos\theta + \beta\sin\theta)} \cos\theta \sin\theta d\theta \approx \frac{e^{\alpha\frac{N}{2}(\frac{\sqrt{2}}{2}-1)}}{2} \sinh\left(\frac{\sqrt{2}N}{4}\beta\right)$$

$$\frac{e^{\alpha\frac{N}{2}(\frac{\sqrt{2}}{2}-1)}}{2} \sinh\left(\frac{\sqrt{2}N}{4}\beta\right) \approx 2p - 1$$

- Resolvendo pelo método das coordenadas polares inversa com o sistema de coordenadas alternativo:

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos \theta + \beta \sin \theta)} \sin \theta \cos \theta d\theta = 2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^1 e^{\frac{N}{2}(\alpha \cos \theta + \beta \sin \theta)} \sin \theta \cos \theta r dr d\theta$$

$$x' = \cos \theta$$

$$y' = \sin \theta$$

$$x'^2 + y'^2 = r^2, \text{ com } r = 1.$$

$$-1 \leq y' \leq 1$$

$$0 \leq x' \leq 1$$

$$\begin{aligned} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos \theta + \beta \sin \theta)} \sin \theta \cos \theta d\theta &= 2 \int_0^1 \int_{-1}^1 e^{\frac{N}{2}(\alpha x' + \beta y')} y' x' dy' dx' \\ 2 \int_0^1 \int_{-1}^1 e^{\frac{N}{2}(\alpha x' + \beta y')} y' x' dy' dx' &= \frac{4}{\beta N} \int_0^1 \left( e^{\frac{N}{2}(\alpha x' + \beta)} + e^{\frac{N}{2}(\alpha x' - \beta)} - \frac{2}{\beta N} e^{\frac{N}{2}(\alpha x' + \beta)} + \frac{2}{\beta N} e^{\frac{N}{2}(\alpha x' - \beta)} \right) x' dx' \\ \frac{4}{\beta N} \int_0^1 \left( e^{\frac{N}{2}(\alpha x' + \beta)} + e^{\frac{N}{2}(\alpha x' - \beta)} \right) x' dx' &= \frac{8}{\alpha \beta N^2} \left( e^{\frac{N}{2}(\alpha + \beta)} + e^{\frac{N}{2}(\alpha - \beta)} \right) \left( 1 - \frac{2}{\alpha N} \left( 1 - e^{-\frac{\alpha N}{2}} \right) \right) \\ \frac{8}{\beta^2 N^2} \int_0^1 \left( e^{\frac{N}{2}(\alpha x' + \beta)} - e^{\frac{N}{2}(\alpha x' - \beta)} \right) x' dx' &= \frac{16}{\alpha \beta^2 N^3} \left( e^{\frac{N}{2}(\alpha + \beta)} - e^{\frac{N}{2}(\alpha - \beta)} \right) \left( 1 - \frac{2}{\alpha N} \left( 1 - e^{-\frac{\alpha N}{2}} \right) \right) \end{aligned}$$

$$\frac{e^{\beta \frac{N}{2}} + e^{-\beta \frac{N}{2}}}{2} = \cosh \left( \beta \frac{N}{2} \right)$$

$$\frac{e^{\beta \frac{N}{2}} - e^{-\beta \frac{N}{2}}}{2} = \sinh \left( \beta \frac{N}{2} \right)$$

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos \theta + \beta \sin \theta)} \sin \theta \cos \theta d\theta = \frac{8e^{\frac{\alpha N}{2}}}{\alpha \beta N^2} \left[ \cosh \left( \beta \frac{N}{2} \right) - \frac{2}{\beta N} \sinh \left( \beta \frac{N}{2} \right) \right] \left( 1 - \frac{2}{\alpha N} \left( 1 - e^{-\frac{\alpha N}{2}} \right) \right)$$

$$\frac{Ne^{\beta \frac{N}{2}}}{2k} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos \theta + \beta \sin \theta)} \cos \theta \sin \theta d\theta = \frac{4e^{\frac{N}{2}(\alpha + \beta)}}{\alpha \beta N k} \left[ \cosh \left( \beta \frac{N}{2} \right) - \frac{2}{\beta N} \sinh \left( \beta \frac{N}{2} \right) \right] \left( 1 - \frac{2}{\alpha N} \left( 1 - e^{-\frac{\alpha N}{2}} \right) \right)$$

$$\frac{4e^{\frac{N}{2}(\alpha + \beta)}}{\alpha \beta N k} \left[ \cosh \left( \beta \frac{N}{2} \right) - \frac{2}{\beta N} \sinh \left( \beta \frac{N}{2} \right) \right] \left( 1 - \frac{2}{\alpha N} \left( 1 - e^{-\frac{\alpha N}{2}} \right) \right) = \cotgh \left( \beta \frac{N}{2} \right) - \frac{2}{\beta N}$$

$$\cotgh \left( \beta \frac{N}{2} \right) - \frac{2}{\beta N} = 2p - 1$$



### Apêndice C: Obtendo as relações entre $\alpha$ e $\beta$ através do cálculo da variância de $P(x)$

$$Var(x) = E(x^2) - [E(x)]^2 = pN(1 - p)$$

$$Var(x) = \int_0^N x^2 \frac{e^{\alpha\sqrt{x(N-x)} + \beta x}}{k} dx - p^2 N^2$$

Utilizando a mesma substituição trigonométrica dos Apêndices A e B:

$$x - \frac{N}{2} = \frac{N}{2} \sin \theta$$

$$x = \frac{N}{2} (1 + \sin \theta)$$

$$dx = \frac{N}{2} \cos \theta d\theta$$

$$0 \leq x \leq N \text{ e } -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}.$$

$$\frac{1}{k} \int_0^N x^2 e^{\alpha\sqrt{x(N-x)} + \beta x} dx = \frac{N^3}{8k} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (1 + \sin \theta)^2 e^{\frac{N}{2}(\alpha \cos \theta + \beta \sin \theta + \beta)} \cos \theta d\theta$$

$$(1 + \sin \theta)^2 = 1 + 2\sin \theta + \sin^2 \theta$$

Dividindo a integral resultante em partes:

$$\frac{N}{2} e^{\beta \frac{N}{2}} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos \theta + \beta \sin \theta)} \cos \theta d\theta = k$$

$$\frac{N}{2k} e^{\beta \frac{N}{2}} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos \theta + \beta \sin \theta)} \sin \theta \cos \theta d\theta = 2p - 1$$

Com isso obtém-se:

$$Var(x) = \frac{N^2}{4} + \frac{N^2}{2} (2p - 1) + \frac{N^3}{8k} e^{\beta \frac{N}{2}} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos \theta + \beta \sin \theta)} \sin^2 \theta \cos \theta d\theta - p^2 N^2$$

A integral restante será calculada por:

- Método de Simpson composto, pelo mesmo motivo pelo qual os cálculos que envolveram a esperança utilizaram tal método:

$$n = 4$$

$$f(\theta) = e^{\frac{N}{2}(\alpha \cos \theta + \beta \sin \theta)} \cos \theta \sin^2 \theta$$

$$h = \frac{\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)}{4} = \frac{\pi}{4}$$

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f(\theta) d\theta \approx \frac{\pi}{12} \left[ f\left(-\frac{\pi}{2}\right) + 4f\left(-\frac{\pi}{4}\right) + 2f(0) + 4f\left(\frac{\pi}{4}\right) + f\left(\frac{\pi}{2}\right) \right]$$

$$f\left(-\frac{\pi}{2}\right) = f(0) = f\left(\frac{\pi}{2}\right) = 0$$

$$f\left(-\frac{\pi}{4}\right) = \frac{\sqrt{2} e^{\frac{\sqrt{2}N}{4}(\alpha - \beta)}}{4}$$

$$f\left(\frac{\pi}{4}\right) = \frac{\sqrt{2} e^{\frac{\sqrt{2}N}{4}(\alpha + \beta)}}{4}$$

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f(\theta) d\theta \approx \frac{\sqrt{2}\pi}{6} e^{\frac{\sqrt{2}N}{4}\alpha} \left[ \frac{e^{\frac{\sqrt{2}N}{4}\beta} + e^{-\frac{\sqrt{2}N}{4}\beta}}{2} \right]$$

$$\frac{e^{\frac{\sqrt{2}N}{4}\beta} + e^{-\frac{\sqrt{2}N}{4}\beta}}{2} = \cosh\left(\frac{\sqrt{2}N}{4}\beta\right)$$

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f(\theta) d\theta \approx \frac{\sqrt{2}\pi}{6} e^{\frac{\sqrt{2}N}{4}\alpha} \cosh\left(\frac{\sqrt{2}N}{4}\beta\right)$$

$$\frac{N^3 e^{\beta \frac{N}{2}}}{8k} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos \theta + \beta \sin \theta)} \cos \theta \sin^2 \theta d\theta \approx \frac{\sqrt{2}N^2 e^{\alpha \frac{N}{2}(\frac{\sqrt{2}}{2}-1)}}{16} \cosh\left(\frac{\sqrt{2}N}{4}\beta\right)$$

Reunindo e reorganizando as partes integradas é obtido:

$$Var(x) \approx \frac{N^2}{4} + \frac{N^2(2p-1)}{2} + \frac{\sqrt{2}N^2 e^{\alpha \frac{N}{2}(\frac{\sqrt{2}}{2}-1)}}{16} \cosh\left(\frac{\sqrt{2}N}{4}\beta\right) - p^2 N^2$$

$$\frac{N^2}{4} + \frac{N^2(2p-1)}{2} + \frac{\sqrt{2}N^2 e^{\alpha \frac{N}{2}(\frac{\sqrt{2}}{2}-1)}}{16} \cosh\left(\frac{\sqrt{2}N}{4}\beta\right) - p^2 N^2 \approx pN(1-p)$$

$$\frac{\sqrt{2} e^{\alpha \frac{N}{2}(\frac{\sqrt{2}}{2}-1)}}{16} \cosh\left(\frac{\sqrt{2}N}{4}\beta\right) \approx \frac{p(1-p)}{N} + \frac{(2p-1)^2}{16}$$

$$e^{\alpha \frac{N}{2}(\frac{\sqrt{2}}{2}-1)} \approx \frac{16p(1-p) + N(2p-1)^2}{\sqrt{2}N \cosh\left(\frac{\sqrt{2}N}{4}\beta\right)}$$

- Utilizando o método das coordenadas polares inversa:

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos\theta + \beta \sin\theta)} \sin^2\theta \cos\theta d\theta = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos\theta + \beta \sin\theta)} (1 - \cos^2\theta) \cos\theta d\theta$$

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos\theta + \beta \sin\theta)} (1 - \cos^2\theta) \cos\theta d\theta = \frac{2}{N} e^{-\beta \frac{N}{2}} k - \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos\theta + \beta \sin\theta)} \cos^3\theta d\theta$$

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos\theta + \beta \sin\theta)} \cos^3\theta d\theta = 2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^1 e^{\frac{N}{2}(\alpha \cos\theta + \beta \sin\theta)} \cos^3\theta r dr d\theta$$

$$x' = \cos\theta$$

$$y' = \sin\theta$$

$$x'^2 + y'^2 = r^2, \text{ com } r = 1.$$

$$-1 \leq y' \leq 1$$

$$0 \leq x' \leq 1$$

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos\theta + \beta \sin\theta)} \cos^3\theta d\theta = 2 \int_0^1 \int_{-1}^1 e^{\frac{N}{2}(\alpha x' + \beta y')} x'^3 dy' dx'$$

$$\begin{aligned}
2 \int_0^1 \int_{-1}^1 e^{\frac{N}{2}(\alpha x' + \beta y')} x'^3 dy' dx' &= \frac{4}{\beta N} \int_0^1 (e^{\frac{N}{2}(\alpha x' + \beta)} - e^{\frac{N}{2}(\alpha x' - \beta)}) x'^3 dx' \\
\frac{4}{\beta N} \int_0^1 (e^{\frac{N}{2}(\alpha x' + \beta)} - e^{\frac{N}{2}(\alpha x' - \beta)}) x'^3 dx' &= \frac{8}{\alpha \beta N^2} \left( x'^3 - x'^2 \frac{6}{\alpha N} + x' \frac{24}{\alpha^2 N^2} - \frac{48}{\alpha^3 N^3} \right) \left( e^{\frac{N}{2}(\alpha x' + \beta)} - e^{\frac{N}{2}(\alpha x' - \beta)} \right) \Big|_0^1 \\
&= \frac{16 e^{\frac{\alpha N}{2}}}{\alpha \beta N^2} \left[ 1 - \frac{6}{\alpha N} + \frac{24}{\alpha^2 N^2} - \frac{48}{\alpha^3 N^3} \left( 1 - e^{-\frac{\alpha N}{2}} \right) \right] \sinh \left( \beta \frac{N}{2} \right) \\
\frac{N^3}{8k} e^{\beta \frac{N}{2}} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{N}{2}(\alpha \cos \theta + \beta \sin \theta)} \sin^2 \theta \cos \theta d\theta &= \frac{N^2}{4} - \frac{2N e^{\frac{N}{2}(\alpha + \beta)}}{\alpha \beta k} \left[ 1 - \frac{6}{\alpha N} + \frac{24}{\alpha^2 N^2} - \frac{48}{\alpha^3 N^3} \left( 1 - e^{-\frac{\alpha N}{2}} \right) \right] \sinh \left( \beta \frac{N}{2} \right) \\
\frac{2N e^{\frac{N}{2}(\alpha + \beta)}}{\alpha \beta k} \left[ 1 - \frac{6}{\alpha N} + \frac{24}{\alpha^2 N^2} - \frac{48}{\alpha^3 N^3} \left( 1 - e^{-\frac{\alpha N}{2}} \right) \right] \sinh \left( \beta \frac{N}{2} \right) &= \frac{N^2(\alpha N - 6)}{2 \left[ \alpha N - 2 \left( 1 - e^{-\frac{\alpha N}{2}} \right) \right]} + \frac{12}{\alpha^2}
\end{aligned}$$

$$Var(x) = \frac{N^2}{4} + \frac{N^2}{2} (2p - 1) + \frac{N^2}{4} - \frac{N^2(\alpha N - 6)}{2 \left[ \alpha N - 2 \left( 1 - e^{-\frac{\alpha N}{2}} \right) \right]} - \frac{12}{\alpha^2} - p^2 N^2$$

$$Var(x) = pN(1 - p)$$

Reorganizando a expressão acima é obtido:

$$\frac{N^2(\alpha N - 6)}{2 \left[ \alpha N - 2 \left( 1 - e^{-\frac{\alpha N}{2}} \right) \right]} + \frac{12}{\alpha^2} = pN(1 - p)(N - 1)$$

## Apêndice D: Cálculo da Esperança e da Variância da Distribuição Normal

A esperança da distribuição normal é dada por:

$$\begin{aligned}
 E(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
 E(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} (x - \mu + \mu) e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
 E(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} \left( \int_{-\infty}^{+\infty} (x - \mu) e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx + \int_{-\infty}^{+\infty} \mu e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \right) \\
 E(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} \left( \left( -\sigma^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right)_{-\infty}^{+\infty} + \mu \int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \right) \\
 E(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} (0 + \mu \sqrt{2\pi\sigma^2}) \\
 E(x) &= \mu
 \end{aligned}$$

A variância da distribuição normal é calculada da seguinte maneira:

$$\begin{aligned}
 var(x) &= E(x^2) - E^2(x) \\
 var(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} \left( \int_{-\infty}^{+\infty} x^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \right) - \left( \int_{-\infty}^{+\infty} \frac{x e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx \right)^2 \\
 \int_{-\infty}^{+\infty} \frac{x e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx &= \mu \\
 var(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} \left( \int_{-\infty}^{+\infty} (x - \mu) x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx + \int_{-\infty}^{+\infty} \mu x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \right) - \mu^2 \\
 var(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} \left( \int_{-\infty}^{+\infty} (x - \mu) x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \right) + \mu \int_{-\infty}^{+\infty} \frac{x e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx - \mu^2
 \end{aligned}$$

Para resolver a primeira integral, utiliza-se o método de integração por partes:

$$\int u dv = uv - \int v du$$

$$u = x$$

$$dv = (x - \mu)e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$$du = dx$$

$$v = -\sigma^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$var(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \left( \left( -x\sigma^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right)_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} \sigma^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \right) + \mu^2 - \mu^2$$

Por L'hôpital:

$$\lim_{x \rightarrow +\infty} \left( -\sigma^2 \frac{\frac{dx}{dx}}{\frac{de^{-\frac{(x-\mu)^2}{2\sigma^2}}}{dx}} \right) = \lim_{x \rightarrow -\infty} \left( -\sigma^2 \frac{\frac{dx}{dx}}{\frac{de^{-\frac{(x-\mu)^2}{2\sigma^2}}}{dx}} \right) = \lim_{x \rightarrow +\infty} \left( -\sigma^2 \frac{\sigma^2}{(x - \mu)e^{-\frac{(x-\mu)^2}{2\sigma^2}}} \right)$$

$$\lim_{x \rightarrow +\infty} \left( -\sigma^2 \frac{\sigma^2}{(x - \mu)e^{-\frac{(x-\mu)^2}{2\sigma^2}}} \right) = 0$$

$$var(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \left( 0 + \sigma^2 \int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \right) + 0$$

$$var(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \left( \sigma^2 \sqrt{2\pi\sigma^2} \right)$$

$$var(x) = \sigma^2$$

## Apêndice E: Cálculo da Esperança e da Variância da Distribuição Binomial

A esperança da distribuição Binomial é dada por:

$$\begin{aligned}
 E(x) &= \sum_{x=0}^{\infty} xp(x) = \sum_{x=0}^n x \binom{n}{x} p^x (1-p)^{n-x} \\
 &= \sum_{x=0}^n \frac{xn!}{x!(n-x)!} p^x (1-p)^{n-x} \\
 &= \sum_{x=0}^n \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x} \\
 &= np \sum_{x=1}^n \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} (1-p)^{n-x} \\
 &= np \sum_{x=1}^n \binom{n-1}{x-1} p^{x-1} (1-p)^{n-x}
 \end{aligned}$$

Fazendo  $y = x - 1$ :

$$\begin{aligned}
 E(x) &= np \sum_{y=0}^{n-1} \binom{n-1}{y} p^y (1-p)^{n-y-1} \\
 \sum_{y=0}^{n-1} \binom{n-1}{y} p^y (1-p)^{n-y-1} &= 1 \\
 E(x) &= np
 \end{aligned}$$

A variância da distribuição binomial é calculada da seguinte forma:

$$\begin{aligned}
 var(x) &= E(x^2) - E^2(x) \\
 var(x) &= \sum_{x=0}^n x^2 p(x) - n^2 p^2 = \sum_{x=0}^n x^2 \binom{n}{x} p^x (1-p)^{n-x} - n^2 p^2 \\
 var(x) &= \sum_{x=0}^n [x(x-1) + x] \binom{n}{x} p^x (1-p)^{n-x} - n^2 p^2 \\
 var(x) &= \sum_{x=0}^n x(x-1) \binom{n}{x} p^x (1-p)^{n-x} + \sum_{x=0}^n x \binom{n}{x} p^x (1-p)^{n-x} - n^2 p^2
 \end{aligned}$$

$$\sum_{x=0}^n x \binom{n}{x} p^x (1-p)^{n-x} = np$$

$$\text{var}(x) = \sum_{x=0}^n \frac{x(x-1)n!}{x!(n-x)!} p^x (1-p)^{n-x} + np - n^2 p^2$$

$$\text{var}(x) = \sum_{x=0}^n \frac{x(x-1)n!}{x(x-1)(x-2)!(n-x)!} p^x (1-p)^{n-x} + np - n^2 p^2$$

$$\text{var}(x) = \sum_{x=0}^n \frac{n!}{(x-2)!(n-x)!} p^x (1-p)^{n-x} + np - n^2 p^2$$

$$\text{var}(x) = n(n-1)p^2 \sum_{x=2}^n \frac{(n-2)!}{(x-2)!(n-x)!} p^{x-2} (1-p)^{n-x} + np - n^2 p^2$$

Fazendo  $y = x - 2$ :

$$\text{var}(x) = n(n-1)p^2 \sum_{y=0}^{n-2} \binom{n-2}{y} p^y (1-p)^{n-y-2} + np - n^2 p^2$$

$$\sum_{y=0}^{n-2} \binom{n-2}{y} p^y (1-p)^{n-y-2} = 1$$

$$\text{var}(x) = n(n-1)p^2 + np - n^2 p^2 = n^2 p^2 - np^2 + np - n^2 p^2 = np - np^2$$

$$\text{var}(x) = np(1-p)$$



## Apêndice F: Cálculo da Convolução entre a Aproximação Contínua da Distribuição Binomial e a Distribuição Normal

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

$$P(x) = \frac{e^{\alpha\sqrt{x(N-x)} + \beta x}}{k}$$

$$f(\tau) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\tau-\mu)^2}{2\sigma^2}}$$

$$P(x - \tau) = \frac{e^{\alpha\sqrt{(x-\tau)(N-x+\tau)} + \beta(x-\tau)}}{k}$$

A partir da equação 2 obtém-se:

$$f(x) * P(x) = \int_{-\infty}^{\infty} f(\tau)P(x - \tau)d\tau$$

$$f(x) * P(x) = \frac{1}{k\sqrt{2\pi\sigma^2}} \int_{x-N}^x e^{-\frac{(\tau-\mu)^2}{2\sigma^2} + \alpha\sqrt{(x-\tau)(N-x+\tau)} + \beta(x-\tau)} d\tau$$

Para solucionar a integral acima, será feita a seguinte substituição trigonométrica:

$$x - \tau - \frac{N}{2} = \frac{N}{2} \operatorname{sen}\theta$$

$$\tau = x - \frac{N}{2}(1 + \operatorname{sen}\theta)$$

$$dx = \frac{N}{2} \cos\theta d\theta$$

$$x - N \leq \tau \leq x \text{ e } -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}.$$

Fazendo as substituições e as manipulações matemáticas necessárias, chega-se a seguinte relação:

$$f(x) * P(x) = \frac{N e^{-\frac{\beta N \sigma^2 + \left(x - \frac{N}{2}\right)^2 - 2\mu x}}{2k\sqrt{2\pi\sigma^2}} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{-\frac{\alpha N \sigma^2 \cos\theta + N \sin\theta \left(\frac{N}{2} - x + \beta N \sigma^2\right) + \left(\frac{N}{2} \sin\theta + \mu\right)^2}{2\sigma^2}} \cos\theta d\theta$$

A integral resultante é de difícil resolução. Para solucioná-la, será utilizado o método de integração de Simpson:

$$\int_a^b f(\theta) d\theta \approx \frac{b-a}{6} \left[ f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right]$$

$$f(\theta) = e^{-\frac{\alpha N \sigma^2 \cos\theta + N \sin\theta \left(\frac{N}{2} - x + \beta N \sigma^2\right) + \left(\frac{N}{2} \sin\theta + \mu\right)^2}{2\sigma^2}} \cos\theta$$

$$a = -\frac{\pi}{2}$$

$$b = \frac{\pi}{2}$$

$$f(a) = 0$$

$$f(b) = 0$$

$$f\left(\frac{a+b}{2}\right) = e^{-\frac{\alpha N \sigma^2 + \mu^2}{2\sigma^2}}$$

$$\int_a^b f(\theta) d\theta \approx \frac{2\pi}{3} e^{-\frac{\alpha N \sigma^2 + \mu^2}{2\sigma^2}}$$

Com isso, chega-se à seguinte relação para a convolução entre  $P(x)$  e  $f(x)$ :

$$f(x) * P(x) \approx \sqrt{\frac{\pi}{2\sigma^2}} \frac{N e^{-\frac{(\alpha+\beta)N\sigma^2 + \left(x - \frac{N}{2}\right)^2 + (\mu-x)^2 - x^2}}{2\sigma^2}}{3k}$$