



**UNIVERSIDADE FEDERAL DO PARANÁ**  
**DEPARTAMENTO DE ESTATÍSTICA**

**RAZÃO DE PROBABILIDADE BAYESIANA COMO  
EVIDÊNCIA PARA TESTE DE VÍNCULO DE  
PATERNIDADE**

**Curitiba**  
**2012**

**Dianna Lo Wai Yan**

**Maurício Fernandes do Nascimento Junior**

**RAZÃO DE PROBABILIDADE BAYESIANA COMO  
EVIDÊNCIA PARA TESTE DE VÍNCULO DE  
PATERNIDADE**

Projeto apresentado à disciplina de  
Laboratório de Estatística do Curso de  
Estatística do Departamento de Estatística  
do Setor de Ciências Exatas da  
Universidade Federal do Paraná.

Orientadora: Prof<sup>ª</sup>. PhD Silvia Emiko  
Shimakura

Co-orientador: Prof. Dr. Marcelo Malaghini

**Curitiba  
2012**

## Sumário

1. Introdução .....	1
2. Objetivo.....	1
2.1 Objetivo Geral.....	2
2.2 Objetivo Específico.....	2
3. Matérias e Métodos .....	2
3.1 DNA.....	2
3.2 Materiais.....	3
3.3 Razão de probabilidades bayesianas .....	3
4. Cronograma de Execução do Trabalho de Conclusão de Curso. ....	4
5. Referências Bibliográficas .....	5

## 1. Introdução

Na área da ciência forense a diversidade de evidências que comprova de um crime é de suma importância. Muitas vezes, apenas as provas encontradas no local do crime não são suficientes para acusar o indivíduo, isso se deve ao fato delas poderem ser muito comuns. Para apoiar à tomada de decisões as técnicas estatísticas são amplamente utilizadas no auxílio de análise e interpretação de dados. Portanto, os estudos dessas técnicas tornam-se essenciais para os peritos criminais, tendo que recorrer aos métodos estatísticos.

Algumas das fontes de evidência mais comuns são: fragmentos como vidros, tecidos, saliva, sêmen, sangue e entre outros. Sendo que os materiais biológicos na área criminal são utilizados para a realização de teste de DNA. A genética forense pode ser dividida em duas partes, a primeira como prova de um crime e a segunda, umas das aplicações mais conhecidas, teste de vínculo de paternidade.

A primeira aplicação na história moderna de análise forense de DNA foi realizada para resolver um caso de estupro e assassinato na década de 1980 pelo Professor Sir Alec Jeffreys da Universidade de Leicester, Reino Unido. Na época os pesquisadores conseguiram descobrir regiões altamente variáveis do DNA, capazes de individualizar uma pessoa, este caso pioneiro demonstrou o potencial do DNA como uma ferramenta nas investigações forense.

No Brasil, o Instituto de Criminalística do Paraná, referência na análise forense do estado, foi criado em 16 de Maio de 1935. O Laboratório de Genética Molecular Forense do Instituto foi inaugurado em 25 de Abril de 2002, sendo um dos precursores na análise de genética na identificação humana em instituições de perícia oficial no Brasil. No período entre 1998 até 2008, pelo instituto, foram analisados 850 casos criminais, sendo que dos casos 48% eram identificação de cadáveres, 32% de crimes sexuais, 12% paternidade criminais, 5% evidências coletadas em localística e 3% troca de bebês em maternidade.

Neste trabalho, os dados foram fornecidos pelo Instituto de Criminalística de Paraná que serão analisados, posteriormente, no software estatístico R. Como forma de pesquisa, visitas técnicas ao Instituto de Criminalística são realizadas para poder obter um conhecimento maior do trabalho dos especialistas no campo e eventuais resoluções de questionamentos.

## **2. Objetivo**

### 2.1 Objetivo Geral

O objetivo deste trabalho é avaliar os resultados de teste de vínculo de paternidade, usando este como evidência para mensurar a probabilidade da relação entre os indivíduos testados.

### 2.2 Objetivo Específico

Avaliar e quantificar a evidência produzida para trios de paternidade: Mãe, Filho e Suposto Pai. Assim como verificar a influência da probabilidade a priori.

## **3. Matérias e Métodos**

### 3.1 DNA

O DNA é uma molécula relativamente simples. É composto por um núcleo de trifosfato, um açúcar desoxirribose e uma de quatro bases nitrogenadas. As bases nitrogenadas são adenina, guanina, citosina e timina. A informação do DNA está contida em todas as células do corpo. Ele tem o formato de uma hélice dupla nas quais cada bloco de DNA está ligado a um segundo bloco. As combinações dos blocos são obrigatoriamente, uma base adenina com uma base timina e uma base citosina com uma guanina.

Dentro de cada célula do corpo existe aproximadamente 3.200.000.000 pares de base (bp) de informação. Esses pares são divididos em (no caso dos humanos) 23 cromossomos, os humanos possuem 46 cromossomos, 23 herdados do pai e 23 herdados da mãe.

Para a análise forense não é necessário ter o conhecimento de todo o genoma da pessoa em questão. Ao comparar o DNA humano com o DNA do primo mais próximo (os chimpanzés) observa-se que apenas 5% dele é diferente. Dentro do genoma humano as pessoas compartilham uma igualdade de 99,9%.

Dentre todos os alelos disponíveis para a análise, 13 loci foram escolhidos para serem utilizados. Esses 13 loci variam bastante entre indivíduos, são fáceis e baratos de serem obtidos, não estão sobre pressão seletiva, possuem baixo índice de mutação e geram um perfil de simples interpretação e fácil comparação entre laboratórios.

### 3.2 Materiais

Os dados genéticos a serem utilizados neste trabalho foram disponibilizados pelo **Instituto de Criminalística do Paraná** e a frequência genética da população brasileira é obtida pelo artigo “Brazilian population database for the 13 STR loci of the AmpF/STR® Profile Plus™ and Carofiler™ multiplex kits”.

### 3.3 Razão de probabilidades bayesianas

Uma das formas utilizadas para verificar a força de evidência para vínculo entre dois indivíduos é a razão entre as verossimilhanças, a qual é chamada de **PI** (Paternity Index):

$$PI = \frac{P(G_F | G_M, G_{SP}, H_A)}{P(G_F | G_M, G_{SP}, H_D)}$$

Em que:

$G_F$  é o genótipo do filho

$H_A$  é a hipótese da acusação

$G_M$  é o genótipo da mãe

$H_D$  é a hipótese da defesa

$G_{SP}$  é o genótipo do suposto pai

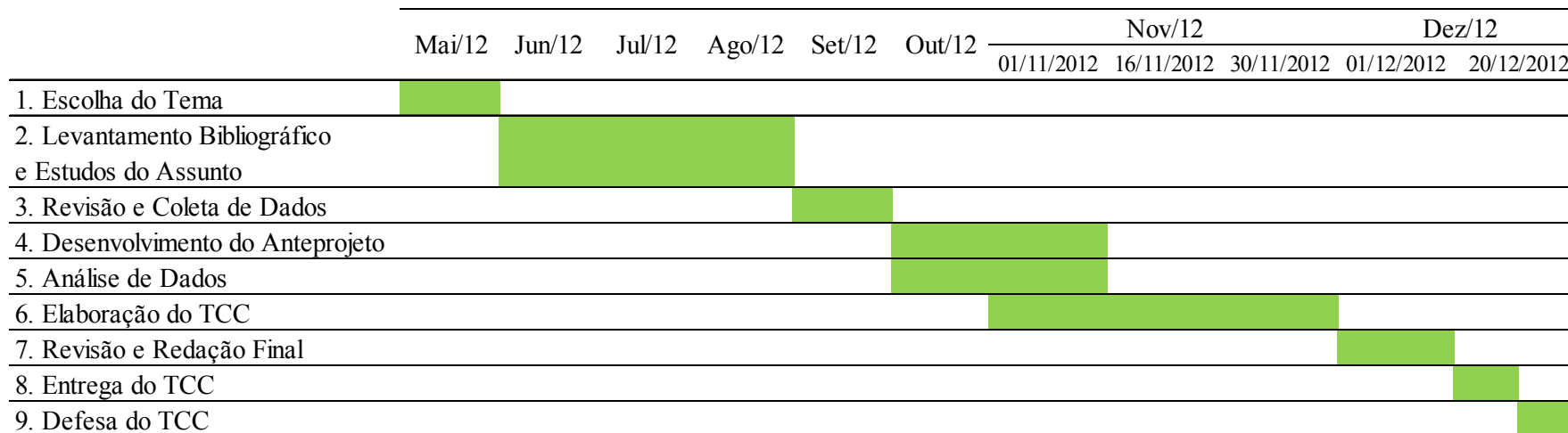
A razão da verossimilhança entre a hipótese da acusação e da defesa representa a força da evidência que a análise de DNA apresenta. Sendo que ambas as verossimilhanças são calculadas usando como base o quadrado de PUNNETT e a frequência alélica geral da população.

Para o cálculo da probabilidade de paternidade precisa-se estipular uma probabilidade a priori, do suposto pai ser o verdadeiro pai da criança, a qual é definida a partir das provas subjetivas como a relação entre a mãe e o suposto pai da criança. A probabilidade de paternidade é definida da seguinte forma:

$$Pp = \frac{PI \times \theta}{(PI \times \theta) \times (1 - \theta)}$$

O objetivo do nosso estudo é através dos dados do Instituto de Criminalística e a utilização de métodos estatísticos, evidência por meio da razão de probabilidade bayesiana, aprimorarem e exemplificarem os casos de teste de vínculo de paternidade.

#### 4. Cronograma de Execução do Trabalho de Conclusão de Curso.



## 5. Referências Bibliográficas

1. Goodwin, W., Linacre, A. and Hadi, S., *An Introduction to Forensic Genetics*, Chichester, 2007
2. Evett, I. W., Weir, B. S., *Interpreting DNA Evidence* Sinauer Associates. ISBN 0-87983-155-4
3. Buckleton, J., Triggs, C. M., *Forensic DNA Evidence Interpretation*, Florida, 2005
4. Butler, J. M., *Fundamentals of Forensic DNA Typing*, San Diego, 2009
5. Lucy, D., *Introduction to Statistics for Forensic Scientists*, Chichester, 2005
6. Grattapaglia, D. Schmidt, A. B., Costa e Silva, C., Stringher, C. Fernandes, A. P. & Ferreira, M.E. 2001. *Brazilian population database for the 13 STR loci of the AmpFlSTR® Profiler Plus™ and Cofiler™ multiplex kits. Forensic Science International* 118(1) 91:94.
7. Instituto de Criminalística. Disponível em:  
< <http://www.ic.pr.gov.br> > Acesso em: 30 Outubro 2012
8. Visitas técnicas ao Instituto de Criminalística do Paraná.