



**UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO**  
**DEPARTAMENTO DE ESTATÍSTICA E INFORMÁTICA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA**

**HÉRBETES DE HOLLANDA CORDEIRO**

---

**MODELOS PROBABILÍSTICOS APLICADOS**  
**À BIOMETRIA**

---

Recife, dezembro de 2005

HÉRBETES DE HOLLANDA CORDEIRO

---

---

**MODELOS PROBABILÍSTICOS APLICADOS  
À BIOMETRIA**

---

Trabalho apresentado ao Programa de Pós-graduação em Biometria do Departamento de Estatística e Informática da Universidade Federal Rural de Pernambuco como requisito parcial para obtenção do grau de Mestre em Biometria.

Orientador: Prof. Dr. Gauss Moutinho Cordeiro

Recife, dezembro de 2005

## **TERMO DE APROVAÇÃO**

HÉRBETES DE HOLLANDA CORDEIRO

### **MODELO PROBABILÍSTICO APLICADO À BIOMETRIA**

Dissertação apresentada como requisito parcial para a obtenção do Título de Mestre em Biometria, do Departamento de Estatística e Informática – Universidade Federal Rural de Pernambuco, pela seguinte banca examinadora:

Prof. Dr. Borko Stosic  
Universidade Federal Rural de Pernambuco

Profa. Dra. Jacira Guiro Carvalho da Rocha  
Universidade Federal de Pernambuco

Prof. Dr. Enivaldo Carvalho da Rocha  
Universidade Federal de Pernambuco

Recife, dezembro de 2005.

## **DEDICATÓRIA**

---

À minha esposa, **Maria das Graças Aguiar Cordeiro**, que sempre me incentivou, acreditando que a ousadia e o erro são caminhos para as grandes realizações.

## **AGRADECIMENTOS**

---

Aos meus filhos: **Cristian** (Doryane), **Hérbetes Jr.**(Cláudia), e **Maria Rachel** (Paulo) pelo incentivo diuturno.

Ao **Programa de Pós-Graduação de Biometria**, da Universidade Federal Rural de Pernambuco, pela oportunidade de cursar este Mestrado.

Ao Professor **Dr. Eufrázio de Souza Santos**, Coordenador do Curso, pelo entusiasmo, dedicação, zelo administrativo e ensinamentos ministrados.

Ao Professor **Dr. Gauss Montinho Cordeiro**, Pesquisador Emérito, Orientador deste trabalho.

Aos Professores **Dr. Borko Stosic** e **Dra. Maria Adélia**, pela amizade fraterna e carinho demonstrados ao longo dessa jornada.

Ao Professor **Dr. Enivaldo Rocha**, da UFPE, pela amizade e parceria na revisão estrutural do Modelo Socioeconômico delineado para a Malária no Brasil.

Aos demais **Professores do Curso de Mestrado**, todo o meu respeito e admiração.

## **Três Coisas**

“De tudo ficaram três coisas:

A certeza de que, estamos sempre começando...

A certeza de que é preciso continuar...

A certeza de que seremos interrompidos antes de terminar...

Portanto, devemos:

Fazer da interrupção, um caminho novo...

Da queda, um passo da dança...

Do medo, uma escada...

Do sonho, uma ponte...

Da procura, um encontro...”

**(Fernando Pessoa)**

O termo “Biometria” significa, literalmente, medida da vida. No mundo da segurança ou criminalidade, refere-se aos métodos automatizados para identificação de pessoas com base em suas características físicas ou aspectos comportamentais. Neste trabalho, procedeu-se uma análise dos principais Sistemas Biométricos de Identificação, de modo a integrá-los às demais áreas de abrangência da Biometria, que incluem desde o diagnóstico e tratamento de doenças, estudos sobre Botânica, Zoologia, Epidemiologia e Bioestatística, até as mais relevantes aplicações contemporâneas, como “Células Tronco” e “Projeto Genoma”. Em outra vertente, realizou-se a descrição de alguns Modelos Probabilísticos e uma pesquisa de suas aplicações à Biometria. Por outro lado, afigura-se também como fulcro desta dissertação, o “Estudo da Malária” e o delineamento dos modelos sócio-econômicos para representá-la no Brasil, onde continua respondendo por grande número de óbitos, principalmente na região Norte. Para materializar este último objetivo, foram coletados em 2004 dados sócio-econômicos e de incidência desta doença no país, em 90 Municípios das regiões Norte, Sul e Sudeste, dados esses estruturadores dos modelos obtidos. Ao final, é feita uma análise dessa infecção nos citados pólos de desenvolvimento, com apresentação de resultados e sugestões para ulteriores trabalhos de modelagem.

**Palavras-chave :** Biometria. Malária. Modelos probabilísticos. Regressão Linear Múltipla

The term "Biometria" means, literally, measured of the life. In the world of the security or crime, one on the basis of mentions the methods to it automatized for identification of people its physical characteristics or mannering aspects. In this work, an analysis was proceeded from the main Systems Biométricos de Identificação, in order to integrate them it the too much areas of abrangência of the Biometria, that include since the diagnosis and treatment of illnesses, studies on Botany, Zoology, Epidemiologia and Bioestatística, until the most excellent applications contemporaries, as "Cells Trunk" and "Genoma Project". In another source, it was become fullfilled basically description of some discrete and continuous-chosen Probabilist Models for being the most cited in Biométrica literature and after an exhausting survey of its respective applications. On the other hand, it is also figured as fulcrum of this dissertação, the "Study of the Malaria" and the delineation of the partner-economic models to represent it in Brazil, where it continues answering for great number of deaths, mainly in the region North. To materialize this last objective, they had been collected in 2004 partner-economic data and of incidence of this illness in the country, in 90 Cities of the regions North, South and Southeast, data these estruturadores of the gotten models. To the end, an analysis of this infection in the cited polar regions of development is made, with presentation of results and suggestions for later works of modeling.

**Key-Words:** Biometria. Malaria. Probabilist models. Multiple Linear regression

## **LISTA DE QUADRO E TABELAS**

---

Quadro 1	Regiões, estados e municípios estudados
Tabela 1	Distribuição do Número de Casos Notificados de Malária X Variáveis Sócio-econômicas por Municípios das Regiões Norte, Sudeste e Sul, Brasil, 2004
Tabela 2	Distribuição das variáveis explanatórias do modelo nacional para os casos de malária no Brasil, 2004
Tabela 3	Distribuição das variáveis explanatórias dos modelos regionais nos casos de malária no Brasil, 2004

## **LISTA DE SIGLAS E ABREVIATURAS**

---

a.C	Antes de Cristo
CIA	Central Intelligence Agence
AIDS	Acquired Immune Deficiency
DATASUS	Departamento de Informática do SUS
d.C	Depois deCristo
DDT	Dicloro Difenil Tricloroetano
DPF	Departamento de Polícia Federal
EUA	Estados Unidos da América
Exp	Exponencial
FBI	Federal Bureau of Investigation
fdp	Função de densidade de probabilidade
IBGE	Instituto Brasileiro de Geografia e Estatística
IC	Intervalo de Confiança
ID	Impressão Digital
INI	Instituto Nacional de Identificação
IPEA	Instituto de Pesquisa Econômica Aplicada
MV	Máxima Verossimilhança
NASA	American Space Agence
OMS	OrganizaçãoMundial de Saúde
STF	Supremo Tribunal Federal

<b>1</b>	<b>INTRODUÇÃO</b> .....	12
<b>2</b>	<b>SISTEMAS BIOMÉTRICOS DE IDENTIFICAÇÃO</b> .....	14
<b>2.1</b>	<b>Revisão da Literatura</b> .....	14
<b>2.2</b>	<b>Tipos de Sistemas Biométricos de Identificação</b> .....	16
<b>2.3</b>	<b>Perspectivas dos avanços biométricos para o futuro</b> .....	22
<b>3</b>	<b>MODELOS PROBABILÍSTICOS APLICADOS À BIOMETRIA</b> .....	25
<b>3.1</b>	<b>Introdução</b> .....	25
<b>3.2</b>	<b>Principais modelos discretos</b> .....	25
<b>3.3</b>	<b>Principais modelos contínuos</b> .....	33
<b>4</b>	<b>ESTUDO DA MALÁRIA</b> .....	46
<b>4.1</b>	<b>Introdução</b> .....	46
<b>4.2</b>	<b>Aspectos epidemiológicos da malária</b> .....	47
<b>4.3</b>	<b>Modelos Probabilísticos históricos da Malária</b> .....	50
<b>4.4</b>	<b>Estruturação Atual de um Modelo Epidemiológico</b> .....	54
<b>4.5</b>	<b>A malária no mundo</b> .....	54
<b>4.6</b>	<b>A malária no Brasil</b> .....	55
<b>5</b>	<b>MATERIAL E MÉTODOS</b> .....	57
<b>6</b>	<b>RESULTADOS</b> .....	62
<b>7</b>	<b>DISCUSSÃO</b> .....	67
<b>8</b>	<b>CONCLUSÕES</b> .....	68
<b>9</b>	<b>SUGESTÃO</b> .....	69
	<b>REFERÊNCIAS BIBLIOGRÁFICAS</b> .....	70
	<b>ANEXOS</b>	

A biometria é uma ciência de relevante importância para a humanidade. As aplicações práticas das tecnologias biométricas são diversas e estão se expandindo magistralmente, conforme as novas necessidades emergentes. Sua classificação está baseada nas características intrínsecas do ser humano, como impressões digitais, geometria das mãos, retina, características faciais e outras como o reconhecimento da voz e a assinatura dinâmica.

Nosso estudo buscou, através da Revisão da Literatura, apresentar uma análise dos principais sistemas biométricos de identificação e ressaltar os avanços da biometria para o futuro, propiciando, desta forma, uma visão sistêmica da sua abrangência, como ciência. O nosso objetivo principal, foi dar ênfase especial às aplicações biométricas, para os principais modelos probabilísticos, buscando suprir a carência desse tipo de informação na literatura pertinente, o que acreditamos ser de grande valia para os futuros pesquisadores.

Historicamente, modelos outros têm sido usados para avaliar as doenças infecto-contagiosas, estabelecendo, desta forma, as bases epidemiológicas para combatê-las.

Ronald Ross (1897) descobriu que a malária era transmitida pelo mosquito *Anopheles* (fêmea), chegando à conclusão, através de seus modelos probabilísticos, adiante abordados, que essa doença se extingiria se a densidade do mosquito (número de mosquitos/ pessoa) fosse inferior a 40.

A malária é uma doença tropical endêmica em muitos países, vitimando anualmente milhões de pessoas em todo o mundo. Cerca de 40% da população mundial vive em áreas com risco de transmissão de malária, resultando em não menos que 300 milhões de pessoas infectadas no mundo a cada ano, mais de 90% em países africanos, com um número de mortes entre 1 e 1,5 milhão. A transmissão ocorre em mais de 100 países da América do Norte (México), América Central, América do Sul (principalmente na Bacia Amazônica), Caribe

(República Dominicana e Haiti), África, Ásia (Subcontinente Indiano, Sudeste Asiático e Oriente Médio), Europa Oriental e Oceania.

No Brasil, a malária é endêmica na Região Norte, que responde por 80% dos casos em todo Território Nacional. Por este motivo, além do estudo da malária, abordando todos os seus aspectos epidemiológicos, fez-se, neste trabalho, o delineamento de modelos socioeconômicos para representá-la no Brasil e nas Regiões que apresentam fortes desigualdades sociais.

## **2. SISTEMAS BIOMÉTRICOS DE IDENTIFICAÇÃO**

---

### 2.1 Revisão da Literatura

A História não é apenas um enumerar de fatos, datas e eventos, mas uma seqüência de acontecimentos. Longe de ser estático, o conhecimento histórico leva o homem a compreender situações existenciais concretas, pretéritas ou atuais, determinadas pelo tempo e espaço.

Nesse contexto, segundo a Enciclopédia Delta Larousse, o primeiro sinal individualizador criado pelo homem para determinar a identidade dos seus semelhantes e das coisas que o rodeavam foi o nome. Porém, esta forma de identificação não teve êxito devido a fatores como adulteração e homonímia. Mais tarde, em meados do século XIX, sentindo a necessidade de identificar e afastar pessoas consideradas nocivas à Sociedade, o homem empregou (em escravos e criminosos) métodos considerados inumanos, como a marcação com ferro incandescente e outros que causavam mutilações físicas.

Nesse afã em busca de segurança, surgiram os métodos de identificação de Anfosso (1896), utilizando o Craniógrafo; o Oftalmostatométrico, de Capdeville (1903) e o da radiografia das mãos, de Tamassia (1908), entre outros.

Com os avanços da ciência da computação (através de softwares e hardwares) os sistemas de identificação biométricos ficaram mais rápidos e seguros, pois se fundamentam nas características intrínsecas do ser humano.

Dentre eles, podemos citar: o reconhecimento da face, da geometria das mãos, da assinatura, da voz, da íris, da retina e das impressões digitais.

Segundo Vucetich (1891), as impressões papilares classificam-se em: digitais (Datiloscopia), palmares (Quiroscopia) e plantares (Pedoscopia). Elas possibilitam a identificação, registro e confrontação de indivíduos, de forma rápida, barata e confiável, constituindo-se no método mais empregado para este fim.

Para Bressan (2002), todos os sistemas de identificação são calcados em três conceitos fundamentais:

- Segurança baseada em algo que o indivíduo possui (como, por exemplo, a chave de um cofre);
- Segurança em algo que o indivíduo sabe (como, por exemplo, uma senha);
- Segurança decorrente de uma característica biométrica (como, por exemplo, sua Impressão Digital – ID).

Assim, teoricamente, qualquer característica humana, seja ela física, comportamental ou fisiológica, pode ser usada para a identificação biométrica, desde que satisfaça às seguintes premissas:

- a) Universalidade: todas as pessoas devem possuir a característica a ser medida.
- b) Singularidade: essa característica não pode ser igual em pessoas diferentes.
- c) Permanência: significa que a característica não pode variar no tempo.
- d) Mensurabilidade: refere-se à precisão da identificação
- e) Aceitabilidade: indica o quanto as pessoas estão dispostas a aceitar os sistemas biométricos.
- f) Proteção: refere-se à facilidade ou dificuldade de enganar o sistema com técnicas fraudulentas.

Hong, em 1998, estabeleceu que os sistemas biométricos são divididos em duas partes: o módulo de registro e o módulo de identificação.

No módulo de registro, as características são obtidas por um leitor biométrico e processadas através de um módulo extrator desses dados, gerando um gabarito que será armazenado. Por sua vez, o módulo de identificação faz a captura das características biométricas no momento do acesso, também gerando uma representação através do módulo

extrator, equivalente à do gabarito armazenado. Finalmente, o módulo de identificação confronta essas informações com o padrão armazenado para estabelecer a identidade. A coincidência entre o padrão armazenado e o coletado em tempo real permitirá que o usuário tenha acesso ao sistema.

#### ❖ Aplicações

Gumz (2002) apresenta aplicações nos procedimentos forenses (identificação criminal, segurança de presídios e identificação de cadáveres), civis (carteira de identidade, carteira de motorista, carteira de saúde e identificação na travessia de fronteiras) e comerciais (comércio eletrônico e serviços bancários, entre outros).

## **2.2 Tipos de Sistemas Biométricos de Identificação**

### 2.2.1 Reconhecimento da face

Para realizar o reconhecimento da face, é necessária uma pequena câmara e um programa para o processamento da imagem. O que o software de reconhecimento faz é registrar pontos estratégicos na face, capazes de definir proporções, distâncias e formas de cada elemento do rosto, como olhos, nariz, queixo, malares, pavilhão auricular, entre outros.

O programa identifica corretamente o rosto do usuário, mesmo que ele tenha mudado o corte ou a cor dos cabelos ou acrescentado, sem exageros, acessórios como óculos, barba ou bigode. É importante ressaltar que alguns sistemas modernos utilizam desde padrões de redes neurais até varreduras infra-vermelhas para detectar esses pontos estratégicos. Outros podem executar testes animados para evitar fraude por fotografia.

#### ❖ Aplicações

- O Departamento de Imigração e Naturalização dos Estados Unidos usa esse método de identificação para tornar mais rápido e eficiente o trânsito regular de cidadãos que trabalham na fronteira do México.
- Em Newham, região metropolitana de Londres, foi implantado, em 1998, um circuito fechado de TV, com mais de 300 câmeras espalhadas em pontos estratégicos e que filmam a rotina da cidade. Em tempo real, as faces das pessoas são comparadas com um banco de dados de criminosos ou suspeitos. Com esta medida, a polícia é alertada sempre que o sistema identifica alguém suspeito. A criminalidade, em decorrência, reduziu cerca de 35%.

#### 2.2.2 Geometria das mãos

Esse estudo vem sendo empregado desde 1970. Baseia-se no fato de que, virtualmente, não existem duas pessoas com mãos idênticas e que o formato das mesmas não sofre mudanças significativas após certa idade. O sistema de identificação utiliza um scanner para capturar e registrar as dimensões da mão, o formato dos dedos, áreas e proporções das articulações, além disso, medidas tridimensionais de pontos selecionados são tomadas e convertidas em um modelo geométrico, sendo posteriormente confrontadas com aquelas existentes no arquivo.

#### ❖ Aplicações

- Nos Jogos Olímpicos de Sidney (1966) a geometria das mãos foi um dos tipos de segurança usados para identificar os atletas e demais participantes do evento.

### 2.2.3 Reconhecimento da retina

O padrão das veias da retina é a característica com a maior garantia de unicidade que uma pessoa pode ter, sendo considerado um dos métodos biométricos de identificação mais seguros. Os abalizadores de retina utilizam um laser de baixa intensidade e uma câmera para a captura de sua imagem, a qual será confrontada com a existente no arquivo. Olhos falsos, lentes de contato e mutações decorrentes de doenças não quebram a segurança do sistema.

#### ❖ Aplicações

- O Federal Bureau of Investigation (FBI), Central Intelligence Agency (CIA) e North American Space Agency (NASA), que exigem altíssimo nível de segurança em suas instalações, usam esse tipo de identificação.

### 2.2.4 Reconhecimento da Íris

Íris é o anel que circunda o olho. Cada íris possui uma estrutura única que forma um padrão complexo, que pode ser usado para identificar um indivíduo. A captura da imagem é feita por uma câmera. Um software compara o padrão de íris capturado com o armazenado no banco de dados, para decidir se foram originados do mesmo olho. Probabilisticamente, o reconhecimento da íris tem uma margem de erro de apenas 0,0008%, enquanto que o de DNA é de aproximadamente 0,05%.

#### ❖ Aplicações

- O Banco Unimed of Texas foi a primeira instituição financeira dos Estados Unidos a implantar o reconhecimento da íris em seus caixas eletrônicos.
- Os aeroportos Charlotte/Douglas, nos Estados Unidos, e Flughafen, na Alemanha, utilizam esse sistema para identificar passageiros, funcionários e agentes de segurança.

### 2.2.5 Reconhecimento da voz

Para esse método é necessário um microfone ou telefone para capturar a voz do usuário. O programa de identificação faz uma análise probabilística dos padrões harmônicos e não somente uma comparação entre reproduções de uma voz. Essa característica é importante para evitar tentativas de fraude, como por exemplo, com uso de gravadores.

Todos os sistemas que abalizam a voz estão amplamente fundamentados na tecnologia de processamento da fala, utilizando o método da análise de Fourier para encontrar os espectros de frequências que assinalam as características da voz, que não são iguais para indivíduos diferentes.

#### ❖ Aplicações

- Em Chicago, nos Estados Unidos, hospitais utilizam o reconhecimento da voz para controlar o acesso de funcionários aos berçários.

### 2.2.6 Reconhecimento da assinatura manuscrita

Existem dois métodos de identificação: o que examina a assinatura já escrita, (comparando-a com a imagem do modelo armazenado) e o que estuda a dinâmica da assinatura, analisando seu ritmo, contato com a superfície, tempo total, pontos de curvas, laços, velocidade e aceleração. Os dispositivos usados para essa análise dinâmica são as canetas ópticas e as superfícies sensíveis.

Como todas as características comportamentais, as assinaturas estão sujeitas ao humor do usuário, ao ambiente e ao tipo de papel, entre outras variáveis. Desta forma, as assinaturas de algumas pessoas são muito consistentes, enquanto que as de outras variam muito.

#### ❖ Aplicações

- Podem ser usados em bancos, cartórios, concursos públicos, entre outros.

#### 2.2.7 Reconhecimento das Impressões Digitais (Datiloscopia)

Datiloscopia é a ciência que trata da identificação de pessoas através da comparação de suas impressões digitais. Oriunda das palavras *Daktylos* (dedos) e *Skopêin* (examinar), representa um método de identificação empregado há mais de um século, sendo considerada a tecnologia biométrica mais prática, segura e econômica que existe, classificando-se em civil, criminal e clínica (INSTITUTO NACIONAL DE IDENTIFICAÇÃO- DEPARTAMENTO DE POLÍCIA FEDERAL, 1987).

Kehdy (1968) e Appes (2004) estabeleceram os seguintes postulados:

- a) Perenidade (universalidade): é a propriedade que tem o desenho digital de existir desde o sexto mês de vida fetal até a putrefação cadavérica, quando se dá o deslocamento da derme. Cada pessoa é classificada por um conjunto de características que permanecem por toda a vida.
- b) Imutabilidade (unicidade): propriedade que tem o desenho digital de não mudar, isto é, cada pessoa tem as suas características próprias.
- c) Variabilidade: propriedade que tem o desenho digital de variar de dedo para dedo e de pessoa para pessoa.
- d) Classificabilidade: propriedade que tem o desenho digital de poder ser classificado para fins de arquivamento e pesquisas.

Para o seu estudo e compreensão, as impressões digitais podem ser divididas em diversas classes, de acordo com a sua topologia geométrica. Eduard Henry propôs a divisão em cinco classes, sendo o seu sistema de classificação (designado de “Henry System”)

baseado no número e localização de pontos chamados de singulares (deltas e linhas do sistema nuclear).

Para Tavares Junior (1990), as impressões digitais compõem-se estruturalmente de: linhas pretas, linhas brancas, deltas, pontos característicos, poros e linhas brancas albedactiloscópicas.

Costa (2000), por sua vez, ressalta que as minúncias (pontos característicos) são acidentes encontrados nas criptas papilares e que têm a finalidade de firmar a unicidade das IDs. Essas minúncias são resumidamente classificadas em duas categorias: aspectos básicos e aspectos compostos. Os aspectos básicos são as cristas finas (ridge ending) e as cristas bifurcadas (bifurcations)

Quanto aos aspectos compostos, formados a partir dos básicos, aparecem as ilhas (Islands), as cristas curtas (Short Ridge), as esporas (Spurs) e os cruzamentos (Crossover). Antheus (2002) considerou que o “núcleo” representava o centro de gravidade de uma ID.

Appol (2004) definiu o delta como sendo o ângulo ou triângulo formado pelas cristas papilares. Para Bresson (2002), os sistemas de identificação através das ID analisam basicamente a posição das minúncias. Todavia, sistemas mais modernos verificam também outras características que aparecem nos dedos tais como: arcos, voltas, posição do(s) delta(s) e núcleos, entre outras com o auxílio de leitores digitais.

No que se refere à identificação de um indivíduo, Kehdy (1968) e Tavares Junior (1991) consideram que duas impressões digitais são consideradas idênticas quando apresentam 12(doze) ou mais pontos característicos com a mesma configuração. Na muitos países esse critério têm validade jurídica para identificação positiva de casos criminais.

#### ❖ Aplicações

- As ID têm sido usadas no controle de acesso a caixas automáticos de bancos e registros de saúde.
- Em projetos-piloto em andamento no País, a ID capturada por sensores ópticos funciona também como senha única na Web, passaporte para dados sigilosos em redes corporativas, inibidor de fraudes no e-commerce e facilitador de controle de recursos humanos.
- No judiciário, as informações sigilosas do Supremo Tribunal Federal (STF) agora estão mais protegidas, graças a um sistema de reconhecimento de ID, que permite (ou não) acesso à rede interna de computadores. Em substituição às senhas, os funcionários precisam apenas tocar o mouse, onde um scanner captura a impressão digital e faz a identificação do usuário.
- O Congresso Nacional também implantou um sistema de identificação equivalente, para registrar a frequência e autenticidade das votações em plenário.
- Em Recife, no Vestibular 2006 das Universidades Federais, foi realizada, pela primeira vez, a identificação dos candidatos através de suas impressões digitais.

### **2.3 Perspectivas dos avanços biométricos para o futuro**

Desde que o fisiologista alemão Theodor Schwann (1939) lançou as bases da Teoria Celular, pesquisadores em todo o mundo sentiram-se instigados com a possibilidade de geração de um organismo adulto completo a partir de apenas uma célula.

Nos últimos anos, pesquisas com células-tronco foram desenvolvidas buscando tratar doenças que afetam milhões de pessoas, mas o entendimento da formação desse organismo remonta ao início do século XX. Foi nessa época que vários embriologistas, como Hans Spemann e Jaques Loeb (1918), começaram a decifrar os segredos das referidas células, ao

demonstrarem que, quando as duas primeiras células de um embrião de anfíbio são separadas, cada uma delas é capaz de gerar um girino normal e que, mesmo para as quatro primeiras divisões celulares, os núcleos dessas células ainda podiam transmitir todas as informações necessárias à formação desse organismo.

O nascimento da ovelha Dolly (1996) mostrou que isto era possível. Pesquisas recentes mostraram que células-tronco embrionárias e adultas apresentam duas características básicas: são indiferenciadas e têm a capacidade de gerar novas células-tronco, bem como outras células com diferentes funções.

Mais adiante a equipe do biólogo Thomson (1998) da Universidade de Wisconsin, tornou este sonho biotecnológico um pouco mais real, quando conseguiu isolar as primeiras células-tronco de embriões humanos. O uso dessas células para o reparo de órgãos e tecidos lesados, abre portas para uma nova era, rica em possibilidades e batizada de medicina regenerativa, a qual segundo alguns pesquisadores, apresenta um potencial revolucionário comparável ao advento da penicilina.

A principal aplicação da terapia de células-tronco seria em doenças crônico-degenerativas, que afetam principalmente pessoas da terceira idade, gerando grande otimismo e esperança para o tratamento de doenças como o mal de Chagas, Alzheimer, Diabetes, entre outras.

No Instituto do Coração de São Paulo são realizadas, com bons resultados, aplicações diretas de células-tronco em pacientes com insuficiência cardíaca causada pela doença de Chagas e/ou hipertensão arterial sistêmica de origem idiopática.. Neste sentido, pesquisadores alemães da Universidade Dusseldorf obtiveram êxito com transplantes de células-tronco retiradas da medula óssea de seres humanos com problemas cardíacos. Nesta pesquisa, realizada com 40 pacientes- onde 20 se submeteram ao transplante e 20 integravam o grupo-controle- observou-se (três meses após o tratamento) que a proporção de sangue bombeado

passou de 55% para 65% e o percentual de tecido cardíaco danificado reduziu de 33% para 14%.

Segundo Kligerman (2002), o desafio para o futuro consiste na seleção destas células e em sua manipulação em laboratório, permitindo a expansão e utilização mais segura dessa tecnologia. Assim, em um futuro próximo, essas células poderão ser modificadas, permitindo a construção de super células, capazes de reconhecer e destruir células tumorais.

Por enquanto, os estudos com células-tronco embrionárias somente são permitidos em modelos animais. Apesar do entusiasmo dos cientistas e esperanças depositadas por uma parcela significativa da população de poder, um dia, beneficiar-se do conhecimento gerado nessa área, ainda são necessárias muitas pesquisas, financiamentos e disposições políticas, éticas e morais, para que seja formado o cenário ideal ao pleno desenvolvimento dessa terapêutica.

### 3. MODELOS PROBABILÍSTICOS APLICADOS À BIOMETRIA

---

#### 3.1 Introdução

Os modelos podem ser considerados como uma representação da realidade em estudo, servindo para simplificar, descrever e interpretar aquilo que se está estudando. Destarte, um modelo probabilístico é uma representação matemática deduzida de um conjunto de premissas com o duplo propósito de estudar os resultados de um experimento aleatório e predizer o seu comportamento futuro, quando este se realiza sob as mesmas condições, dadas inicialmente.

Ele permite conhecer a distribuição de probabilidades que toma a variável aleatória, através de suas funções de probabilidade, densidade, distribuição e geração de momentos, além dos parâmetros de tendência, variabilidade, assimetria e “kurtosis”

Sob a ótica da biometria, formulações matemáticas têm sido historicamente desenvolvidas para descrever o surgimento de doenças infecciosas, para testar conceitos e para fornecer a compreensão clara (insight) na direção do controle de doenças e formulação de diretrizes.

#### 3.2 Principais modelos discretos

##### 3.2.1 Modelo Binomial

A distribuição binomial pode ser definida usando a expressão binomial  $(q + p)^n = \sum_{k=0}^n \binom{n}{k} p^k q^{n-k}$  como a distribuição da variável aleatória X tal que

$$P_r(X = x) = \binom{n}{x} p^x q^{n-x}, \quad x = 0, 1, 2, \dots, n,$$

Onde  $q + p = 1$ ,  $p > 0$ ,  $q > 0$  e  $n$  é um inteiro positivo.

Quando  $n = 1$  a distribuição é conhecida como a distribuição de Bernoulli.

A média, a variância e a função característica são representadas, respectivamente, por  $\mu = np$ ,  $\mu_2 = npq$  e  $\phi_X^{(t)} = (q + pe^{it})^n$ . Considerando os dados  $x_1, x_2, \dots, x_n$  como valores observados de variáveis aleatórias  $X_1, X_2, \dots, X_n$  com distribuição binomial de índice  $m_i$  e parâmetro  $\pi_i$ , respectivamente, temos a partir de (1.1) que a log-verossimilhança de  $\pi_i$  dado  $x$  é escrita da seguinte forma (CORDEIRO, 1974):

$$\ell(\pi; x) = \sum_{i=1}^n \left[ x_i \log \left( \frac{\pi_i}{1 - \pi_i} \right) + m_i \log(1 - \pi_i) \right]$$

#### ❖ Aplicações

Tem grande importância a distribuição de uma variável aleatória  $X$  que representa o número de sucessos em uma seqüência de Provas de Bernoulli, quando não interessa a ordem em que ocorrem. O emprego da distribuição Binomial (DB) aparece em genética porque características biológicas dependem dos genes que ocorrem em pares, o que possibilitou, por exemplo, que Fisher e Mather's (1936) analisassem, em cobaias, dados provenientes de cabelos lisos e cabelos crespos.

Seal (1949) aplicou essa distribuição para efetuar análises em dados de mortalidade, enquanto que Boswell, Ord e Patil (1979) desenvolveram aplicação em ecologia das plantas.

Mais recentemente, Kaplan e Risco (1982) procederam estudos com o objetivo de determinar o número de nucleotídeos que estão na mesma disposição (sítio) em duas seqüências de DNA.

Seber (1982) cita um grande número de aplicações da D.B. à ecologia animal, com ênfase especial para a determinação do tamanho de uma população. Ensaio do tipo dose-resposta (usados na área de toxicologia), também podem ser tratados pela distribuição binomial. Por exemplo, quando um inseticida é aplicado a um determinado número de insetos, eles respondem (morrem) ou não (sobrevivem) à dose aplicada.

Ainda, na área médica, quando a dose de medicamento é administrada a um grupo de pacientes, eles podem melhorar (sucesso) ou não (falha). De uma forma geral, o objetivo desse tipo de experimento tratado pela distribuição  $Y \sim \text{BIN}(m, \pi)$  é modelar a probabilidade de sucesso  $\pi$  com função de variáveis explanatórias e, então, determinar doses efetivas ( $DL_p$ ) que causam mudanças de estado em  $p\%$  de indivíduos.

### 3.2.2 Modelo de Poisson

Diz-se que uma variável aleatória  $X$  tem distribuição de Poisson como parâmetro  $\theta$  se

$$P_r(X = x) = \frac{e^{-\theta} \theta^x}{x!}, \quad x = 0, 1, 2, \dots, \theta > 0$$

A média, a variância e a função característica são dadas, respectivamente, por

$$\mu = \theta, \quad \mu_2 = \theta \quad \text{e} \quad \phi_X^{(t)} = \exp \{ \theta(e^{it} - 1) \}.$$

#### ❖ Aplicações

A distribuição de Poisson tem sido muito utilizada para muitas situações práticas e em problemas nos quais a probabilidade de ocorrência de um evento é pequena, sendo, por esta razão, comumente designada como distribuição dos eventos raros.

Doob (1953) usou-a para distribuições moleculares e Parzen (1962), para processos de nascimentos.

Haight (1967) faz referências para o seu emprego em Agricultura, Ecologia, Biologia, Medicina, Sociologia e Demografia.

Cliff e Ord (1981) a utilizaram na mensuração de grupos populacionais por unidade de área ou volume.

Frienberg (1982) a empregou para determinar a força ou atividade biológica de uma substância pela comparação de seus efeitos com uma preparação padrão. Nesse experimento, é feita a contagem de colônias de bactérias ou vírus em uma amostra, para diferentes diluições ou condições experimentais quando a contagem direta não é possível, mas a presença ou

ausência do organismo (em sub-amostras) pode ser detectada. O modelo de Poisson é aplicado ainda, segundo Frienberg, para o estudo de incidência de câncer e mortalidade infantil.

Taylor e Karlin (1984) evidenciaram o seu uso nas áreas de Biociência, Medicina, Teoria do Risco e Demografia.

Soong (1986) faz referência à contagem de bactérias numa Placa de Petri, em distribuição de fertilizante (espalhado por um avião) sobre um campo e dispersão de poluentes industriais numa região.

Margolin *et al* (1986) aplicam o modelo de Poisson para determinar mudanças cromatídicas nos estudos de DNA.

### 3.2.3 Modelo Binomial Negativo

Muitos modelos diferentes dão origem à distribuição binomial negativa e, conseqüentemente, há uma variação de definições na literatura. Formalmente, a distribuição binomial negativa pode ser definida em termos da expansão da expressão binomial negativa  $(Q - P)^{-k}$ , onde  $Q = 1 + P$ ,  $P > 0$  e  $K$  é um número positivo real: desta forma a distribuição binomial negativa com parâmetros  $K, P$ , é a distribuição da variável aleatória  $X$  tal que:

$$P_r[X = x] = \binom{k + x - 1}{k - 1} \left(\frac{P}{Q}\right)^x \left(1 - \frac{P}{Q}\right)^k, \quad x = 0, 1, 2, \dots$$

Onde  $Q = 1 + P$ ,  $P > 0$ , e  $K > 0$ .

A média, variância e a função característica são dadas, respectivamente, por

$$\mu = KP, \quad \mu_2 = KP(1 + P) \quad \text{e} \quad \phi_X^{(t)} = (1 + P - pe^{it})^{-K}$$

#### ❖ Aplicações

A distribuição binomial negativa tem se tornado cada vez mais popular como uma alternativa mais flexível para a distribuição de Poisson, especialmente quando há dúvidas

quanto aos requisitos estritos, particularmente o critério de independência, para a distribuição de Poisson ser satisfeita.

Greenwood e Yule (1920), Arbous e Kerrich (1951) e Kemp (1970) têm aplicado essa distribuição em estatísticas de acidentes, enquanto que Furry (1937) e Kendall (1949) mostraram sua aplicação em processos de nascimento e morte.

Sichel (1951) enfatizou o seu uso para análise de dados psicológicos. Alguns anos após, aplicações médicas e militares foram ressaltadas por Bennet e Birch (1964).

Rao (1973) usou essa distribuição para modelar tamanho de população, enquanto Biins (1986) e Perry (1984) a empregaram em exames de dados entomológicos.

### 3.2.4 Modelo Logarítmico

Uma variável aleatória  $X$  tem distribuição logarítmica se

$$\begin{aligned} P_r(X = x) &= \frac{a\theta^x}{x}, \quad x = 1, 2, \dots \\ &= \frac{(x-1)\theta}{x} P_r(X = x-1), \quad x = 2, 3, \dots \end{aligned}$$

onde  $0 < \theta < 1$  e  $a = -[\log(1 - \theta)]^{-1}$ .

A média e a variância são respectivamente:

$$\mu = \frac{a\theta}{1 - \theta} \quad \text{e} \quad \mu_2 = \frac{a\theta(1 - a\theta)}{(1 - \theta)^2} = \mu[(1 - \theta)^{-1} - \mu],$$

onde  $a = \frac{-1}{\log(1-\theta)} = \{\log(1 - \theta)^{-1}\}^{-1}$ .

A função característica é dada pela seguinte expressão:

$$\phi_X^{(t)} = \frac{\log(1 - \theta e^{it})}{\log(1 - \theta)}$$

## ❖ Aplicações

Em Fisher, Corbet e Williams (1943) as distribuições com as séries logarítmicas foram aplicadas com o objetivo de obter amostras de borboletas e também de mariposas capturadas através de armadilhas luminosas.

Williams (1944), usou esse tipo de série para analisar dados de publicações de entomologistas e, no ano seguinte, para o levantamento de outros numerosos conjuntos de dados ecológicos como, por exemplo, a distribuição de parasitas (piolhos) por colônia, numa cabeça.

Witherby (1941), por sua vez, elaborou estudos sobre a distribuição de ninhos de várias espécies de pássaros britânicos.

### 3.2.5 Modelo Hipergeométrico

Considere a situação clássica em que a distribuição Hipergeométrica aparece. Suponha que uma urna contenha  $N$  bolas, onde  $Np$  são brancas e  $N-Np$  são pretas. Se uma amostra aleatória de  $n$  bolas for retirada dessa urna com reposição, então a probabilidade de sair uma bola branca será constante ( $= p$ ) e a distribuição do número  $X$  de bolas brancas retiradas será binomial com parâmetros  $n$  e  $p$ . Contudo, se uma amostra aleatória de  $n$  bolas é retirada dessa urna, sem reposição, então a probabilidade que o número de bolas brancas seja  $X$  é dada por:

$$P[X = x] = \sum \frac{n!}{r_0! r_1! \dots r_s!} p_0^{r_0} p_1^{r_1} \dots p_s^{r_s}, \quad x = 0, 1, 2, \dots, n_s$$

onde  $\max(0, n - N + Np) \leq x \leq \min(n, Np)$ . Esta expressão define uma distribuição hipergeométrica com parâmetros  $n, Np, N$ , todos positivos. A média e a variância são

$$\mu = E(X) = np \quad \text{e} \quad \mu_2 = \text{Var}(X) = \frac{np(1-p)(N-n)}{N-1}.$$

## ❖ Aplicações

Chapman (1952) e Peterson (1986) aplicaram a distribuição Hipergeométrica para estimação de população de animais com os dados obtidos através do processo “captura-recaptura”. Exemplificando: considere que se deseja estimar o número de peixes existentes num pequeno lago. Inicialmente, uma amostra é retirada e um número conhecido “ $m$ ” de peixes é marcado e, em seguida, recolocado no referido lago. Pouco tempo depois (longo o suficiente para haja uma dispersão randômica dos peixes marcados, mas não obstante (longo) para que ocorram mudanças naturais que possam afetar o tamanho da amostra), uma segunda amostra de tamanho “ $n$ ” é retirada do lago e o número “ $x$ ” de peixes marcados é contado.

As assertivas de que as amostras têm tamanhos fixos (“ $m$ ” e “ $n$ ”) e que os peixes se distribuem no lago de forma independente e aleatória, conduzem à utilização de uma distribuição hipergeométrica para  $X$ . Há outras variações envolvendo o modelo básico hipergeométrico e o processo “captura-recaptura”. Por exemplo, a estimação de população marcada  $N$  em estudos epidemiológicos, quando se pretendem determinar o número de anencéfalos nascidos no Hospital Lying-In, em Boston, por várias décadas, pode ser obtido pela contagem do número de casos registrados em duas listas de tamanho “ $n$ ” e “ $m$ ”. Para testar se dicotomias são independentes ou não, Cochran (1954) recomendou a aplicação do “teste exato de Fisher” em substituição ao do qui-quadrado ( $\chi^2$ ).

Ainda Cox e Snell (1989) apoiaram-se nesse tipo de distribuição para analisar dados médicos dicotomizados, como fumantes/não-fumantes e pacientes com câncer/controle.

Em um estudo, procedeu-se à análise de uma amostra de mulheres grávidas fumantes, sendo essas amostras dicotomizadas para recém-nascidos de “baixo peso” ou de “peso normal”. Verificou-se que as gestantes tabagistas apresentaram crescimento intra-uterino retardado em comparação com as gestantes não tabagistas.

### 3.2.6 Modelo de Série Logarítmica

A variável aleatória  $X$  tem uma distribuição de série logarítmica se

$$P_r[X = k] = \frac{\alpha \theta^k}{k} \quad (k = 1, 2, \dots, 0 < \theta < 1).$$

Onde  $\alpha = -[\log(1 - \theta)]^{-1}$

As probabilidades são os termos na expressão da série de  $-\alpha \log(1 - \theta)$ .

A função geratriz de probabilidade é:

$$\alpha \sum_{i=1}^{\infty} (\theta^i / i) t^i = [\log(1 - \theta t)] / [\log(1 - \theta)]$$

#### ❖ Aplicações

Blackman (1935) utilizou esse modelo e obteve um número médio de diferentes espécies de plantas encontradas em áreas quadráticas de tamanhos variados, numa formação de pasto. Fisher (1942) usou sistematicamente essa distribuição em um trabalho publicado nessa época, mostrando a relação entre o número de espécies e o número de indivíduos obtidos em uma amostra aleatória de uma população animal.

Desta forma, essas séries foram utilizadas para obter dados com amostras de borboletas e mariposas capturadas por armadilhas luminosas. Em suas pesquisas, Rowe (1942) relacionou o número de mosquitos capturados (também por armadilhas luminosas) em várias cidades em Iowa, com o número de espécies observadas.

Williams (1947) o aplicou para problemas biológicos e Kendal (1948) para representar o crescimento de populações de tumores.

### 3.3 Principais modelos contínuos

#### 3.3.1 Modelo Gama

Uma variável aleatória  $X$  tem distribuição Gama se sua função de densidade de probabilidade é da forma

$$p_X^{(x)} = \frac{(x - \gamma)^{\alpha-1} \exp[-(x - \gamma)/\beta]}{\beta^\alpha \Gamma(\alpha)}, \quad \alpha > 0, \beta > 0; x > \gamma.$$

Sendo  $\Gamma(\alpha)$  a função gama, definida por:  $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$ . Ela depende de três parâmetros  $\alpha$ ,  $\beta$  e  $\gamma$ . Se  $\gamma = 0$ , a distribuição é considerada uma distribuição GAMA de dois parâmetros, denotada Gama ( $\alpha$ ,  $\beta$ ).

A forma padronizada da distribuição é obtida fazendo  $\beta = 1$  e  $\gamma = 0$ .

Então:

$$p_X^{(x)} = \frac{x^{\alpha-1} e^{-x}}{\Gamma(\alpha)}, \quad x \geq 0.$$

Se  $\alpha = 1$ , obtém-se a distribuição exponencial

$$p_X^{(x)} = e^{-x}, \quad x \geq 0.$$

Se  $\alpha$  é um inteiro positivo, obtém-se a Distribuição de Erlang. Se  $\alpha = \frac{\nu}{2}$ ,  $\beta = 2$  e  $\gamma = 0$  então obtém-se a distribuição de  $\chi^2$  (Qui-Quadrado).

A função característica é dada pela expressão:

$$\phi_X^{(x)} = (1 - it)^{-\alpha}$$

## ❖ Aplicações

A precipitação pluvial é o elemento regulador na agricultura, sendo que a quantidade de chuva, assim como a sua distribuição em certa localidade, pode determinar o tipo de atividade agrícola a ser desenvolvida. Essas informações fornecem subsídios para determinar períodos críticos de escassez de chuvas em uma região, bem como a implantação de culturas que se adaptem a esse tipo de sazonalidade de precipitação atmosférica.

A distribuição Gama tem sido usada para fazer ajustes realísticos para distribuições exponenciais na representação do tempo de vida, bem como na análise de precipitação de chuvas (DAS, 1955; KOTZ e NEUMANN, 1963).

Friedman e Janes (1957) estudaram o problema de distribuição de probabilidade da ocorrência de chuvas e evidenciaram o fato de que são necessárias dados de (no mínimo) 30 anos para que o tamanho da amostra seja representativo.

Salem e Mount (1974) realizaram um estudo comparativo entre a distribuição Gama e a Log-normal, tendo demonstrando que a primeira forneceu uma melhor opção para a compilação de dados de chuvas nos EUA, no período de 1960 a 1969.

Frizzone (1979) estimou a precipitação de chuvas para a região de Viçosa, MG, em períodos de 5, 10, 15 e 30 dias, utilizando cinco modelos probabilísticos. Os estudos demonstraram que a distribuição Gama incompleta se ajustava bem aos dados para períodos de 5 a 30 dias, na referida região.

Castro Neto e Silveira (1983) estudaram a precipitação pluvial provável em períodos de 15 e 10 dias, respectivamente, através da distribuição Gama, tendo concluído que seria possível a utilização de irrigação suplementar na região, o que diminuiria o custo de instalação e operação do sistema de irrigação no município de Lavras, MG.

Dennis e Patil (1984) discutiram a aplicação da distribuição Gama para estatísticas ecológicas. Galote (1987) estudou o ajuste da distribuição Gama para os dados de

precipitação pluvial no município de Belém-PA, utilizando dados coletados no período de 1953 a 1986, agrupados em 10 períodos. Verificou que a distribuição Gama ajustava-se bem aos dados em razão da *hipótese nula* não ser rejeitada em 97,5% dos períodos estudados, o que comprovaria a acurácia desse modelo.

### 3.3.2 Modelo Normal Inverso (Wald)

Suponha que uma partícula se move ao longo de uma linha com uma velocidade uniforme  $v$ . Suponha também que a partícula está submetida a um movimento Brauniano linear. Pode-se demonstrar que o tempo  $X$  requerido para se cobrir uma distância  $d$  é uma variável aleatória com função de densidade de probabilidade ( $fdp$ )

$$p_X^{(x)} = \frac{1}{\sqrt{2\pi\beta x^3}} d \exp \left[ -\frac{(d - vx)}{2\beta x} \right], \quad x > 0.$$

Onde:  $\beta$  é a constante de difusão (SCHRÖDINGER, 1915). Substituindo  $v = \frac{d}{\lambda}$  e  $\beta = \frac{d^2}{\lambda}$  em (1.1), obtém-se a forma padronizada (canônica) para o modelo Gaussiano Inverso de 2 parâmetros:

$$p_X(x; \mu, \lambda) = \left[ \frac{\lambda}{2\pi x^3} \right]^{1/2} \exp \left\{ -\frac{\lambda}{2\mu^2 x} (x - \mu)^2 \right\}, \quad x > 0.$$

A média e a variância de  $X$  são dados, respectivamente, por:  $E(X) = \mu$  e  $\text{Var}(X) = \frac{\mu^3}{\lambda}$

Função Geratriz de Momentos:

$$\phi_X^{(t)} = \exp \left\{ \lambda \mu^{-1} \left[ 1 - \left( 1 - \frac{2\mu^2 t}{\lambda} \right)^{1/2} \right] \right\}$$

Função Geratriz de Cumulantes:

$$R_X^{(t; \phi, \lambda)} = \exp \left\{ \lambda \mu^{-1} \left[ 1 - \left( 1 - \frac{2\mu^2 t}{\lambda} \right)^{1/2} \right] \right\}$$

Outras características importantes, nomeadas por (CORDEIRO, 2004) são dadas por:

- a) Coeficiente de assimetria

$$3\sqrt{\mu/\lambda}$$

- b) Coeficiente de Curtose

$$3 + 15\mu/\lambda$$

- c) Moda

$$\mu \left\{ \left( \frac{1 + 9\mu^2}{4\lambda^2} \right)^{1/2} - \left( \frac{3\mu}{2\lambda} \right) \right\}$$

#### ❖ Aplicações

Segundo Wasan (1968), a distribuição Gaussiana têm sido usada para explicar o deslocamento de partículas influenciadas pelo “Movimento Brauniano”, sendo particularmente aplicada para estudar a movimentação de partículas em suspensão coloidal submetida a um campo elétrico (TWEEDIE, 1947).

Não obstante o seu usual emprego em “Análise Seqüencial”, alguns Autores tem usado esta distribuição para várias aplicações práticas conectadas com processos de difusão em condições limites. Recentemente, a distribuição Gaussiana Inversa tem sido aplicada em dados de reabilitação e de análise de tempo de vida (PADGETT e TSAI, 1986), em testes acelerados de vida (BHATTCHARYYA e FRIES, 1982) e em ajuste nas distribuições de tempo, especialmente nos casos de preponderância de falha (CHHIKARA e FOLKS, 1977).

Cordeiro (2004) mostra outras aplicações que incluem modelagem de variáveis, como: “Tempo de Sobrevivência” e “Tempo gasto para injetar uma substância no Sistema Biológico”.

### 3.3.3 Modelo Exponencial

A variável aleatória  $X$  tem uma distribuição exponencial (ou exponencial negativa) de parâmetro  $\alpha$  se ela tem função de densidade de probabilidade dada por:

$$p_X^{(x)} = \alpha^{-1} \exp \left[ -\frac{x - \theta}{\alpha} \right], \quad x > \theta; \alpha > 0.$$

É um caso especial da distribuição gama. Com muita frequência torna-se  $\theta = 0$  e obtém-se a chamada distribuição exponencial uniparamétrica.  $p_X^{(x)} = \alpha^{-1} \exp(-\frac{x}{\alpha})$  com  $\theta = 0$  e  $\alpha = 1$ ,

tem-se a chamada distribuição exponencial padronizada, com função de densidade de probabilidade  $p_X^{(x)} = \exp(-x)$ ,  $x > 0$ .

A função característica é dada pela expressão:

$$\phi_X^{(t)} = (1 - i\alpha t)^{-1} e^{it\theta} \text{ se } X_1, X_2, \dots, X_n \text{ são}$$

variáveis aleatórias independentes cada uma delas com f.d.p dada por (1.2) então os estimadores de máxima verossimilhança de  $\theta$  e  $\alpha$  são:

$$\hat{\theta} = \min(X_1, X_2, \dots, X_n),$$

$$\hat{\alpha} = n^{-1} \sum_{i=1}^n (X_i - \hat{\theta}) = \bar{X} - \hat{\theta}.$$

Se  $\theta$  é conhecido então  $\hat{\alpha} = \bar{X} - \theta$ .

A função de Distribuição ou função acumulada de probabilidades é:

$$F_X^{(x)} = 1 - e^{-\alpha x}, \quad x \geq 0$$

Características numéricas da distribuição

$$\text{Média: } \mu = E(X) = \frac{1}{\alpha}$$

$$\text{Variância: } \sigma^2 = \text{Var}(X) = \frac{1}{\alpha^2}$$

$$\text{Desvio Padrão: } \sigma = \frac{1}{\alpha}.$$

A comprovação de que a distribuição exponencial tem a propriedade de “falta de memória” é imediata.

Se  $X \sim \varepsilon(\sigma)$  então

$$\begin{aligned} p(X > t + s / X > t) &= \frac{p(X > t + s)}{p(X > t)} = \frac{\int_{t+s}^{\infty} \sigma e^{-\sigma x} dx}{\int_t^{\infty} \sigma e^{-\sigma x} dx} \\ &= \frac{e^{-\sigma(t+s)}}{e^{-\sigma t}} = e^{-\sigma s} = p(X > s). \end{aligned}$$

#### ❖ Aplicações

O tempo de vida pode ser representado por uma variável aleatória exponencial.

#### 3.3.4 Modelo Log-normal

Galton (1879) não descarta que, se  $X_1, X_2, \dots, X_n$  são variáveis independentes e positivas e  $T_n = \prod_{i=1}^n X_i$ , então  $\log T_n = \sum_{i=1}^n \log X_i$ , e se as variáveis aleatórias independentes  $\log X_i$  são tais que o teorema do limite central possa ser aplicado, então a distribuição padronizada de  $\log T_n$  tenderá para a distribuição normal quando  $n$  tender para o infinito. Então, nossas condições, a distribuição de  $T_n$  será lognormal.

Desta forma, se há um número  $\theta$  tal que  $\alpha = \log(X - \theta)$  seja normalmente distribuído, a distribuição de  $X$  é dita lognormal. Para este caso, é fundamental que  $X$  tome valores maiores que  $\theta$ , que tenha probabilidade zero de assumir valores abaixo dele.

A distribuição de  $X$  pode ser definida pela equação (1):  $U = \gamma + \delta \log(X - \theta)$  onde  $U$  é uma variável normal e  $\gamma, \delta$  e  $\theta$  são parâmetros. A função de densidade de probabilidade  $X$  é (2):

$$p_X^{(x)} = \delta [(x - \theta)\sqrt{2\pi}]^{-1} \exp \left\{ -\frac{1}{2} [\gamma + \delta \log(x - \theta)]^2 \right\}, \quad x > \theta.$$

Uma notação alternativa troca  $\gamma$  e  $\delta$ , respectivamente, pelo valor esperado ( $\varepsilon$ ) e pelo desvio padrão ( $\sigma$ ) de  $Z = \log(X - \theta)$ .

Estes conjuntos de parâmetros estão assim relacionados

$$\xi = -\frac{\gamma}{\delta} \quad \text{e} \quad \sigma = \delta^{-1}.$$

As expressões (1) e (2), após substituição, ficam:

$$U = \frac{\log(X - \theta) - \xi}{\sigma}$$

$$p_X^{(x)} = [(x - \theta)\sqrt{2\pi} \cdot \sigma]^{-1} \exp \left\{ -\frac{1}{2} \frac{[\log(x - \theta) - \xi]^2}{\sigma^2} \right\}, \quad x > \theta.$$

#### ❖ Aplicações

Oldham (1965) aplicou a distribuição Log-normal para determinar a duração de “ausência de doença” e o “tempo de consulta médica”. Hermanson e Johnson (1967) mostraram que essa distribuição se ajustaria adequadamente às correntes líquidas ou fluxos de inundação.

Essa distribuição tem sido frequentemente encontrada em textos acadêmicos por ser uma boa opção para a distribuição de Weibull, na representação do tempo de vida útil dos produtos manufaturados. É também, no mínimo, uma alternativa mais realística que a Distribuição Normal para representar variáveis como peso, altura e densidade.

#### 3.3.5 Modelo de Weibull

Uma variável aleatória  $X$  tem distribuição Weibull se há valores dos parâmetros  $c(> 0)$ ,  $\alpha(> 0)$  e  $\varepsilon_0$  tais que  $Y = \left(\frac{X - \varepsilon_0}{\alpha}\right)^c$  tem distribuição exponencial padronizada com função de densidade de probabilidade

$$p_Y^{(y)} = e^{-y}, \quad y > 0.$$

A função de densidade de probabilidade Weibull para a variável aleatória  $X$  é:

$$p_X^{(x)} = \frac{c}{\alpha} \left( \frac{x - \varepsilon_0}{\alpha} \right)^{c-1} e^{-\left(\frac{x - \varepsilon_0}{\alpha}\right)^c}, \quad x > \varepsilon_0$$

Função de distribuição acumulada

$$F_X^{(x)} = 1 - e^{-\left(\frac{x-\varepsilon_0}{\alpha}\right)^c}, \quad x > \varepsilon_0$$

Função de sobrevivência

$$R_X^{(x)} = 1 - F_X^{(x)} = e^{-\left(\frac{x-\varepsilon_0}{\alpha}\right)^c}, \quad x > \varepsilon_0$$

Função de risco

$$h_X^{(x)} = \frac{p_X^{(x)}}{R_X^{(x)}} = \frac{c}{\alpha} \left( \frac{x - \varepsilon_0}{\alpha} \right)^{c-1}, \quad x > \varepsilon_0.$$

Esta função é decrescente para  $c > 1$ , constante para  $c = 1$  e crescente para  $c < 1$ . A esperança de vida é obtida de (2):

$$E(X) = \int_x^\infty \frac{R_X^{(u)}}{R_X^{(x)}} du = e^{-\left(\frac{x-\varepsilon_0}{\alpha}\right)^c} \int_x^\infty e^{-\left(\frac{u-\varepsilon_0}{\alpha}\right)^c} du.$$

#### ❖ Aplicações

O fato da distribuição Weibull ser decrescente, constante e crescente, para  $c < 1$ ,  $c = 1$  e  $c > 1$ , respectivamente, tem tornado essa distribuição extremamente útil como modelo de tempo de vida. Apesar da advertência feita por Gorski (1968) sobre a “Weibull Euphoria”, numerosos trabalhos têm aparecido desde 1970, explorando variadas e interessantes aplicações com fulcro nesse modelo, conforme se depreende da relação abaixo:

- a) Formulação de experimentos de carcinogênese (BERRY, 1975).
- b) Análise da relação: pressão sistólica do sangue, colesterol total e fumo, para 14 anos de óbitos em pessoas de uma companhia de Gás, em Chicago (DYER, 1975).
- c) Incidência de câncer pulmonar em fumantes (WHITTEMORE e ALTSCHULER, 1976).
- d) Análise de dados hidro-meteorológicos, meteorológicos e farmacocinéticos (WONG, 1977; VODA, 1978 ; SCHWENKE, 1987), respectivamente.

- e) Pesquisa genética (RINK *et al*, 1979).
- f) Análise de velocidade dos ventos (VAN DER AUWERA, DE MAYER e MALET, 1980); (BARROS e STEVAN, 1983); (DIXON e SWIFT, 1984); (CONRADSEN, NIELSEN e PRAHM, 1984); (TULLER e BRETT, 1984).
- g) Força dos ventos em turbinas geradores de energia (CARLIN e HASLLET, 1982).
- h) Análise de dados de inundação (ZHANG, 1982); (BOES, 1989); (NATHAN e McMAHON, 1990).
- i) Análise de dados de transplantes cardíacos, em Stanford (AITKINH, LAIRD e FRANCIS, 1983).
- j) Modelo dose-resposta para descrever os efeitos do ozônio (RAWLINGS e CURE , 1985).
- k) Análise da curva de sobrevivência de pacientes com câncer após tratamento (CHEN *et al*, 1985).
- l) Distribuição e análise de sistemas migratórios (KAMAROGLOU, LIAW e PAPAGEORGIOU, 1986).
- m) No estudo da distribuição diamétrica do *Ecalyptus grandis* (GLADE, 1986).
- n) Análise de dados de tempestades (SCHUETTE, SALKA e ISRAELSON, 1987).
- o) Análise discriminatória de recorrência do alcoolismo entre veteranos de guerras (HOMAN, PARDO e TRUDEAU, 1987).
- p) Avaliação do efeito da temperatura na germinação cumulativa da alfafa (BAHLER, HILL e BYERS, 1989).

### 3.3.6 Modelo Beta

A distribuição beta  $\beta(p, q)$  é definida pela seguinte função de densidade de probabilidade:

$$f_Y(y; p, q) = \frac{y^{p-1}(1-y)^{q-1}}{B(p, q)}, \quad \text{para } 0 \leq y \leq 1$$

com  $p, q > 0$ .  $p$  e  $q$  são parâmetros de forma a serem estimados. Nossa expressão  $B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$  é a função beta.

É uma f.d.p unimodal para  $p, q > 1$  e não unimodal para  $p, q < 1$  ou  $\frac{p-1}{q-1} \leq 0$ .

As estimativas de M.V. de  $p$  e  $q$  satisfazem o sistema

$$\psi(\hat{p}) - p(\hat{p} + \hat{q}) = \log \bar{y}$$

$$\psi(\hat{q}) - \psi(\hat{p} - \hat{q}) = \frac{1}{n} \sum \log(1 - y_i)$$

Que deve ser resolvido iterativamente pelo método escore.

#### ❖ Aplicações

Janardan e Padmanabhan (1986) usaram a distribuição Beta para modelar variáveis hidrológicas. Harrop e Williams (1989) desenvolveram estudos de porosidade/proporção de vazios de solos, enquanto que Wiley, Herschokoru e Padian (1989) desenharam um modelo para estimar a probabilidade de transmissão do vírus HIV durante o contato sexual entre um infectado e um indivíduo susceptível.

Mais adiante, McNally (1990) empregou essa distribuição no estudo de variáveis que afetam a reprodutividade de vacas e Koppes e Grover (1992) utilizaram-na para determinar o tamanho de progênie em *Escherchia Coli*.

### 3.3.7 Modelo de Laplace

É dado pela seguinte função de densidade de probabilidade

$$p_X^{(x)} = \frac{1}{2\phi} e^{-\frac{|x-\theta|}{\phi}}, \quad -\infty < x < \infty, \quad \phi > 0.$$

Essa distribuição é também chamada de “exponencial dupla”.

A forma da função de densidade de probabilidade padronizada é obtida fazendo  $\theta = 0$  e  $\phi = 1$ , então  $p_X^{(x)} = \frac{1}{2}e^{-|x|}$ .

A função característica é:

$$E[e^{itX}] = (1 + t^2)^{-1}$$

#### ❖ Aplicações

Okubo e Narita (1980) utilizaram esse modelo na predição de tempestades no Japão.

### 3.3.8 Valor Extremo

A Função de densidade de probabilidade é dada por

$$P_r[X \leq x] = \exp \left\{ -e^{-\frac{(x-\varepsilon)}{\theta}} \right\}$$

onde  $\varepsilon$ ,  $\theta(> 0)$  e  $k(> 0)$  são parâmetros. A distribuição correspondente de  $(-X)$  é chamada de distribuição de valor extremo.

Essa distribuição é algumas vezes chamada de “distribuição exponencial dupla” em virtude do seu formato, mas deve-se ter o cuidado de não tomá-la pela distribuição de Laplace, que tem a mesma designação.

O termo “valor extremo” é associado a essa distribuição porque ele pode ser obtido como o limite da distribuição (quando  $n \rightarrow \infty$ ) do maior valor entre  $n$  variáveis aleatórias independentes, cada uma tendo a mesma distribuição contínua. Pela substituição de  $X$  por  $(-X)$ , a distribuição limite do menor valor é obtida. Desta forma, pode-se enfatizar que a teoria

de valores extremos está preocupada com a análise estatística de eventos raros ou extremos, ou seja, com a análise de valores máximos ou mínimos de uma variável aleatória, estatísticas de ordem e valores que excedem um certo limite. Com a sua utilização, pode-se estimar probabilidades (com razoável precisão) nos limites encontrados dentro da amostra e mesmo além deles.

Assim, modelar valores extremos significa ajustar apenas a cauda da distribuição dos dados originais, ressaltando o caráter pouco informativo que se pode ter no interior da distribuição para as ocorrências de baixíssima frequência.

#### ❖ Aplicações

- a) O primeiro trabalho que descreve uma aplicação da distribuição valor extremo para fluxos de inundação foi elaborado por FULLER (1914).
- b) Estudos sobre o fenômeno de ruptura e inundação de maciços de terra (GRIFFITH, 1920).
- c) Modelagem de emissões radioativas e tempo de vida humano, delineamento de descargas de córregos e para estimar o nível de ocorrência desses eventos (GUMBEL, 1937).
- d) Modelagem do tempo de sobrevivência de microorganismos (VELZ, 1947).
- e) Análise da magnitude e das frequências de inundações na bacia do Rio Columbia, medidos pelo U.S. Geological Survey (RANTZ e RIGGS, 1949).
- f) Estudos de dados pluviométricos e de “*run-off*” em pequenas bacias hidrográficas (POTTER, 1949).
- g) Determinação da frequência e velocidade máxima dos ventos (THOM, 1954).
- h) Modelagem de valores anuais, mínimos e máximos, de alguns elementos meteorológicos (JENKINSON, 1955).

- i) Análise de dados de reabilitação (KING, 1959; CANFIELD e BORGMAN, 1975).
- j) Análise estatística de dados de tempestades (SIMIU; FILLIBEN, 1975 e 1976 ; SIMIU; BIETRTRY e s FILLIBEN, 1978).
- k) Predição de inundações (SHEN; BRYSON e OCHOA, 1980).
- l) Predição de resultados com cavalos de raça (HENERY, 1984).
- m) Análise de risco sísmico e terremoto no Pacífico (BURTON e MAKROPOULOS, 1985).
- n) Análise da frequência de inundações (HALL, 1992).
- o) Estimar a distribuição de aerossol e estudos de problemas ambientais (HOPKE e PAATERO, 1993).
- p) Membranas sintéticas (RAJAN, 1993).

### 4.1 Introdução

A malária é uma doença que acomete o homem desde a pré-História.

Originado-se do continente africano, considerado o “Berço da Humanidade”, acompanhou a saga migratória do ser humano pelas regiões do Mediterrâneo, Mesopotâmia, Índia e Sudoeste Asiático. A sua chegada ao novo mundo ainda hoje é motivo de especulações, já que os pesquisadores não dispõem de informações confiáveis, pois existem referências sobre febres sazonais e intermitentes em textos religiosos e médicos bastante antigos, entre assírios, chineses e italianos, vinculando-as à punição de deuses e presença de maus espíritos.

No Século V, a.C., na Grécia, Hipócrates foi o primeiro médico a descartar esta superstição ao associar essas febres às estações do ano e aos locais freqüentados pelos doentes. Depois dele, no Século II, d.C., médicos gregos e romanos desenvolveram trabalhos sobre uma febre conhecida como “Febre Romana” que também ocorria em epidemias cíclicas na Grécia, Itália e em diversas partes da Europa.

Durante quase 1500 anos pouco foi acrescentado ao conhecimento sobre essa doença e seu tratamento. No Século XVII, ainda, padres jesuítas observaram que indígenas da América do Sul usavam a casca de uma árvore nativa chamada “Cinchona” para o tratar alguns tipos de febres. Essa prática se espalhou rapidamente pela Europa e o produto usado, cujo princípio ativo era o “quinino”, passou a ser conhecido como “pó dos jesuítas”.

Finalmente, no Século XVIII, a doença recebeu o nome italiano de “mal aire”, que significa “mau ar” ou “ar insalubre”, pois acreditava-se ser causada por emanações e miasmas provenientes dos pântanos.

Somente no final do Século XIX é que o conhecimento sobre a malária passou por um período de importantes descobertas.

Em 1880, Charles Alphonse Laveran, trabalhando na Argélia, foi o primeiro médico a observar e descrever parasitas da malária no interior de glóbulos vermelhos humanos. Mais adiante, o epidemiologista Ronald Ross (1880), trabalhando na Índia, elucidou o modo de transmissão da doença, ao encontrar formas do parasita da malária no interior de um mosquito que havia se alimentado com sangue de um indivíduo infectado.

O ciclo de desenvolvimento desse Plasmódio no homem e na fêmea do mosquito *Anopheles* foi obtido posteriormente, graças aos pesquisadores italianos Giuseppe Bastianelli(1898) e Batista Grassi(1899) e o seu detalhamento será feito a seguir.

## **4.2 Aspectos epidemiológicos da malária**

### 4.2.1 Agentes Etiológicos:

- a) O Plasmódio Vivax
- b) O Plasmódio Falciparum
- c) O Plasmódio Malariae
- d) O Plasmódio Ovale

### 4.2.2 Ciclo de vida dos parasitas

Nas quatro espécies de plasmódios acima nomeadas, o ciclo de vida é essencialmente o mesmo: apresenta uma fase sexuada exógena (Esporogonia) com multiplicação dos parasitas em certos mosquitos do gênero *Anopheles* e uma fase assexuada endógena (Esquizogonia) com multiplicação no hospedeiro humano.

### 4.2.3 Ciclo de vida do parasita no mosquito

Enquanto os Anofelinos machos se alimentam somente de néctar e seiva vegetal, as fêmeas necessitam de sangue em sua alimentação, para o amadurecimento de seus ovos e

possibilitar a ovoposição. Assim, após ingerir sangue de um hospedeiro humano, contendo as formas sexuadas do parasita (gametócitos), inicia-se uma fase sexuada no interior de seu estômago, com a fecundação e formação de um ovo ou zigoto.

Posteriormente, o zigoto migra através da camada única de células do estômago do mosquito, posicionando-se entre esta e sua membrana basal. Desta forma, por esporogonia, surgem centenas de formas infectantes (esporozoítas) que migram para as glândulas salivares do inseto, as quais poderão ser inoculadas no ser humano, no momento da picada.

#### 4.2.4 Ciclo de vida do parasita no homem

Ao picar um animal ou homem os mosquitos injetam uma pequena quantidade de saliva que serve basicamente de anticoagulante. É nesta saliva que, caso o mosquito esteja infectado, podem ser encontrados os esporozoítas, os quais, após a inoculação, circulam livres pelo sangue durante aproximadamente 30 minutos. Neste período alguns deles são fagocitados e outros podem alcançar o fígado e, no interior das células hepáticas, eles passam por uma primeira divisão assexuada (esquizogonia tecidual).

Decorridos alguns dias, com o surgimento de milhares de parasitas, a célula do fígado se rompe e os plasmódios tem acesso ao sangue, onde invadem os glóbulos vermelhos. Novamente se multiplicam de forma assexuada (esquizogonia eritrocítica), em ciclos variáveis (de 24 a 72 horas), cada parasita produzindo de 8 a 32 novos exemplares, em média.

#### 4.2.5 Transmissor (vetor)

O inseto transmissor da malária é um Culicídeo do gênero *Anopheles*, também conhecido por pernilongo, muriçoca, carapanã e mosquito prego, existindo mais de 400 espécies desse gênero no mundo.

O Período de incubação dura em média 15 dias (desde a picada do infectante até o início dos sintomas da doença). As fêmeas das diversas espécies de Anofelinos colocam os seus ovos em criadouros que podem ser: grandes lagos ou lagoas, remansos de rios e córregos, represas artificiais, valas de irrigação, alagados, manguezais e até mesmo plantas da família das bromeliáceas. Os ovos medem cerca de 0,5 mm de comprimento e são dotados de flutuadores nas laterais, permitindo que permaneçam na superfície da água. Dos ovos surgem as larvas, que se transformam em pupas e depois em adultos, desenvolvimento este que ocorre entre 7 e 20 dias, em função da temperatura ambiental. As fêmeas são fecundadas uma só vez mas podem permanecer colocando ovos férteis entre 2 e 4 dias, durante toda a sua vida, que é de aproximadamente 30 dias.

Quanto à transmissão da doença, o plasmódio chega ao ser humano geralmente através da picada de uma fêmea Anofelina infectada. Como os Plasmódios estão presentes na circulação sanguínea, durante a infecção, a transmissão da malária também pode ocorrer a partir da transfusão de sangue, de transplantes de órgãos, da utilização compartilhada de seringas por usuários de drogas endovenosas ou da gestante para o filho (malária congênita) antes ou durante o parto, e, ainda, em decorrência de acidentes de trabalho com pessoal de laboratório ou hospital.

#### 4.2.6 Aspectos Clínicos

De uma maneira geral todo ser humano é susceptível à malária, mesmo àqueles vitimados diversas vezes, pois a imunidade induzida pela presença do parasita no organismo nunca chega a conferir proteção total.

Por outro lado, o conjunto de sintomas e sinais conhecidos como acesso malárico, caracteriza-se por intenso calafrio, seguido de elevação rápida da temperatura corpórea, acompanhada de náuseas e/ou vômitos, dor de cabeça, dores musculares e abdominais. À

medida que a temperatura começa a baixar, o doente apresenta intensa sudorese, podendo este acesso durar vários minutos ou horas e repetir-se em intervalos diferentes, de acordo com a espécie de Plasmódio. Anemia, falta de apetite, aumento do tamanho do fígado e do baço, fraqueza e distúrbios gastrintestinais também podem ser observados.

A malária (quando detectada correta e precocemente) é uma doença de tratamento bastante simples: para cada espécie de Plasmódio, é utilizado um medicamento (ou associações de medicamentos) em dosagem adequada à situação de cada doente. Todavia, quando há retardo no diagnóstico e a infecção for ocasionada pelo Plasmódio Falciparum, grandes complicações poderão surgir, inclusive a morte do paciente.

Quanto à prevenção, apesar das pesquisas que vem sendo realizadas há muitos anos, não existe uma vacina que confira proteção satisfatória contra a malária. Entretanto, para se obter razoável proteção, deverão ser tomadas medidas de ordem pessoal como: utilização de repelentes químicos e antimaláricos, uso de mosquiteiros sobre as camas, afixação de telas em portas e janelas das habitações e evitar a permanência ao ar livre nos crepúsculos matutino e vespertino.

#### **4.3 Modelos Probabilísticos históricos da Malária**

O médico Inglês Ronald Ross (1897) efetuou pesquisas sobre o ciclo de vida dos parasitas da Malária, protozoários que se multiplicam nos glóbulos vermelhos do sangue do homem e que são transmitidos pelo mosquito do gênero Anopheles.

Por ter revelado os mistérios da malária e enriquecido a ciência com esta descoberta de grande interesse médico-biológico, ele recebeu, em 1902, o 2º Prêmio Nobel de Medicina.

#### 4.3.1 O Modelo Multiplicativo

A abordagem epidemiológica inicial adotada por Ross para explicar a dinâmica de transmissão desta doença foi o “método à posteriori”, iniciando com as estatísticas observadas, procurando ajustá-las às leis analíticas e regredindo até a causa fundamental.

Posteriormente, com a evolução dos estudos e pesquisas efetuados em epidemiologia matemática, ele manifestou sua clara preferência pelo método “à priori” assumindo o conhecimento da causa, modelando a equação diferencial pertinente e finalmente, testando os resultados dos cálculos, comparando-os com as estatísticas observadas.

Ele formulou inicialmente, a “teoria dos acontecimentos dependentes” e, num trabalho sobre “Controle da Malária na Maurîtânia” empregou o seguinte modelo multiplicativo:

$$N = p.m.i.a. b.s.f$$

onde:

$N \rightarrow$  N<sup>o</sup> de novas infecções por mês

$p \rightarrow$  população média

$m \rightarrow$  proporção da população infectada

$i \rightarrow$  proporção infecciosa entre infectados

$a \rightarrow$  média do n<sup>o</sup> de mosquitos/por pessoa

$b \rightarrow$  proporção de mosquitos não infectados que se alimentam no homem

$s \rightarrow$  proporção de mosquitos que sobrevivem durante o período de incubação

$f \rightarrow$  proporção de mosquitos infectados que se alimentam no homem.

Ross chamou de “Densidade Crítica do Mosquito” o valor de  $a$ , dado pela expressão:

$$a = N/ p.m.i.b.s.f$$

Ele considerou, ainda, as seguintes premissas:

- a)  $N^{\circ}$  de recuperações/mês =  $N^{\circ}$  de novas infecções/mês ou seja:  $r.p.m = N$
- b) Proporção de mosquitos não infectados = Proporção de mosquitos infectados, isto é:  $b = f$
- c)  $r =$  Proporção de recuperação.

Probabilidades :  $r = 0,20$     $b = 0,25$     $i = 0,25$     $s = 0,33$

Substituiu essas igualdades em para obter:

$$a = r.p.m / p.m.i.b.s.f. = r / b2.i.s$$

Baseando-se na sua experiência e conhecimentos bioestatísticos e matemáticos, estimou probabilidades subjetivas para as variáveis explicativas do modelo acima, considerando-as adequadas à sua pesquisa epidemiológica:

$$r = 0.20; \quad b = 0.25; \quad i = 0.25 \quad e \quad s = 0.33.$$

Logo, ele obteve para a densidade crítica do mosquito:

$$a = 0,20 / 0,25.0,25.0,25.0,33 = 40$$

Nestes termos, Ross concluiu que a “Malária se extinguiria se a densidade do mosquito pudesse ser reduzida para menos de 40 por pessoa, na população”.

Esta forma de pensar, entretanto, contradizia o critério prevalente na medicina e na saúde, que considera a “eliminação completa do mosquito” como condição “*sine qua non*” para o controle da malária.

#### 4.3.2 O Modelo Exponencial

Ross, mais adiante, deu um tratamento matemático mais sofisticado ao seu modelo malária, usando as seguintes equações diferenciais:

$$\frac{dx}{dt} = -\beta xy \text{ (Velocidade de ocorrência de casos susceptíveis, mês a mês)}$$

$$\frac{dy}{dt} = -\beta xy - \gamma y \text{ (Velocidade de casos infecciosos, mês a mês)}$$

onde:

$x(t) \rightarrow$  n° de casos susceptíveis no instante  $t$

$y(t) \rightarrow$  n° de casos infecciosos no instante  $t$

$\beta \rightarrow$  índice de transmissão

$\gamma \rightarrow$  índice de recuperação

Então:

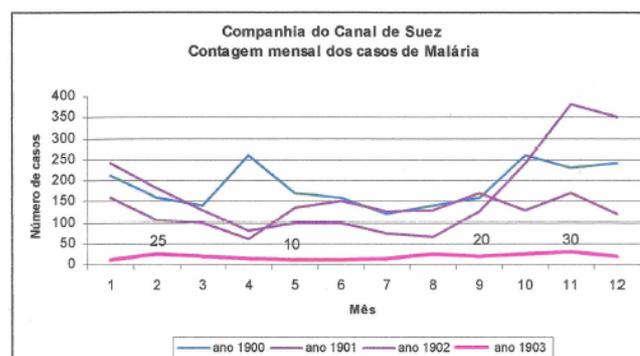
$$\frac{dx}{x} = -\beta y dt \text{ e, integrando, } \log x = -\beta y t \text{ vem } x = e^{-\beta y t}$$

$$\frac{dy}{y} = (\beta x - \gamma) dt \text{ e, integrando, } \log y = (\beta x - \gamma) t \Rightarrow y = e^{(\beta x - \gamma) t}$$

Ao final do seu estudo Ross concluiu que tanto o número de casos susceptíveis ( $x$ ) quanto o número de casos infecciosos ( $y$ ) seguem o modelo exponencial.

#### 4.3.3 Controle da Infecção na Cia do Canal de Suez

Ross testou o seu modelo estatístico na Mauritània (Índia) no período de 1900 a 1903, provando que a doença estava efetivamente sob controle, conforme se depreende no gráfico abaixo:



Fonte: Rosse 1903

#### **4.4 Estruturação atual de um Modelo Epidemiológico**

Os modelos probabilísticos adotados por Ross evoluíram e hoje existe uma especialidade chamada Biomatemática, que tem o computador como ferramenta essencial para realizar modelagens bem mais sofisticadas, enriquecendo as pesquisas em todos os campos do conhecimento.

Assim, utilizando a malária como exemplo, os modelos matemáticos atuais comparam as eficácias dos instrumentos de controle da doença como: o tratamento de infectados, o controle biológico dos vetores, o desenvolvimento de drogas que atuam nas diversas etapas do ciclo de vida dos parasitas, as diferentes estratégias de vacinas e até o uso de mosquitos modificados geneticamente, para se tornarem resistentes às infecções ou serem usados com antígenos vacinais contra essa enfermidade.

Mais recentemente, os modelos tem procurado abordar os aspectos moleculares relacionados à resposta imunológica e a alta taxa de mutação dos protozoários levando-os a padrões evolutivos característicos.

#### **4.5 A Malária no mundo**

Dados da Organização Mundial de Saúde (OMS) indicam que a malária é a doença infecto-contagiosa tropical que mais causa problemas sociais e econômicos no mundo, somente superada pelo número de mortes por AIDS.

Ela é endêmica em 53 países da África, 21 países das Américas, 4 países da Europa e 14 na Região Leste do Mediterrâneo e no Sudeste Asiático. Desta forma, a cada ano ocorrem 350 a 500 milhões de casos, com cerca de 1 milhão de óbitos, sendo 80% deles na África, onde 70% são de crianças com menos de 5 anos, o que equivale a 1 morte a cada 30 segundos.

É importante frisar que a aquisição de malária não é uniforme dentro de uma mesma região, além de sofrer com as estações do ano e ao longo do tempo. Por volta de 1970, os

programas de erradicação da malária, da OMS, livraram dessa doença cerca de 53% da população residente em áreas malarígenas, evitando milhões de mortes e contribuindo para o desenvolvimento socioeconômico de grandes áreas especialmente na Ásia, sul e sudeste da Europa e nas Américas.

#### **4.6 A Malária no Brasil**

Durante a primeira metade do século XX muitas pesquisas foram dedicadas ao controle da malária, especialmente no sentido de reduzir ou eliminar a presença de criadouros do inseto transmissor, o que se mostrou eficaz em algumas situações.

Na década de 1940, governantes de vários países resolveram reagir contra essa enfermidade que matava milhões de pessoas todos os anos. Iniciaram, então, um programa para combater o mosquito *Anopheles* (com o uso do inseticida DDT) e tratar toda a população com um medicamento antimalárico que acabava de surgir, a cloroquina.

Essas ações conjuntas fizeram com que o referido programa fosse um sucesso. No Brasil, por exemplo, os 6 milhões de casos registrados anualmente no citado período, caíram para cerca de 40 mil por ano, na década de 1960.

O problema cresceu a partir de 1970, em decorrência dos Projetos de Desenvolvimento da Amazônia, com a implantação de rodovias, a construção de hidroelétricas, garimpos, atividades extrativistas (acarretando alterações ambientais significativas), e implantação de agrovilas. Além disso, as crises econômicas, a redução das atividades de controle da doença, a elevação dos preços dos inseticidas, a resistência dos Anofelinos aos inseticidas e dos parasitas aos antimaláricos, fizeram a malária expandir-se magistralmente nessa Região.

Em 1999, o país registrou 640 mil casos de malária, com 203 óbitos. Em 2001, esse número caiu para cerca de 390 mil, com 142 óbitos, em virtude de medidas de controle do

vetor, acesso ao diagnóstico laboratorial e ao tratamento eficaz e imediato dos casos suspeitos, verificando-se, então, uma redução de 69 % no número de internações e mortes em hospitais.

Na região da Amazônia Legal, estão situados os focos de maior risco de transmissão da doença (cerca de 500 mil casos, anualmente).

Nos demais estados a doença geralmente ocorre quando é transmitida por pessoas infectadas oriundas dessas áreas malarígenas.

## 5. MATERIAL E MÉTODOS

---

No sentido de obter-se uma visão sistêmica das áreas de abrangência da Biometria, procedeu-se uma análise dos principais sistemas biométricos de identificação, que se baseiam fundamentalmente, nas características intrínsecas do ser humano. Neste sentido, para cada tipo de sistema aventado, foram coletadas informações na literatura especializada.

No que se refere aos modelos publicados, após a pesquisa, foram levantadas as suas correspondentes aplicações à Biometria.

No decorrer do estudo realizado para o delineamento dos modelos socioeconômicos para a Malária no Brasil, constatou-se que a mesma está restrita quase que integralmente à região da Amazônia Legal (formada pelos Estados da região Norte e pelos Estados do Maranhão e Mato Grosso), sendo que nas demais Unidades da Federação, o risco de transmissão local (autóctone) é pequeno ou inexistente.

Nesse contexto, considerando que o mosquito *Anopheles* (fêmea) é o vetor de transmissão dessa enfermidade, evidenciou-se, nesta dissertação, a oportunidade de serem tecidas considerações sobre os modelos probabilísticos (históricos) adotados por Ronald Ross, basicamente no que concerne a determinação da “densidade crítica do mosquito”, ( $a = 40$  mosquitos por pessoa) valor limite para a malária manter-se sob controle.

No período de 1980 a 2003, conforme se depreende das informações veiculadas pelo Ministério da Saúde, resumidas na tabela 1 (Anexo), verifica-se que a malária é endêmica em toda a Região Norte do Brasil, onde as desigualdades sócio-econômicas são mais acentuadas, respondendo por cerca de 80 % dos casos registrados em todo o país. Em contrapartida a infecção está praticamente sob controle nas regiões mais desenvolvidas, como Sul e Sudeste, ocorrendo geralmente essa doença quando é transmitida por pessoas infectadas que vêm de um dos Estados da supracitada região (Norte) ou de outros países, principalmente da África.

Em nosso estudo foram escolhidos (aleatoriamente) dois Estados por região (Norte, Sul e Sudeste), sendo cada um deles representado por 15 dos seus respectivos municípios, igualmente selecionados, os quais encontram-se discriminados no seguinte Quadro 1:

Quadro 1. Regiões, estados e municípios estudados.

Região	Estados	Municípios
Norte	Pará	Altamira; Belém; Marabá; Santarém, Itaituba; Jacareacanga; Tucuruí; Anajás; Santana; Castanhal; Dom Eliseu; Paragominas; Redenção; Itupiranga; Monte Alegre.
Norte	Amazonas	Boca do Acre; Manaus; Lábria; Apuí; Manacapuru; Careiro; Borba; Tefé; São Gabriel; Urucará; Barcelos; Tabatinga; Novo Airão; Iranduba; Carajás
Sul	Rio Grande do Sul	Porto Alegre; Caxias do Sul; Passo Fundo; Frederico Westphalen; Estrela; Maximiliano de Almeida; Torres; Ronda Alta; Alegria; Bento Gonçalves; Lajeado; Ijuí; Taquara.; Sapucaia do Sul; São Leopoldo
Sul	Paraná	Curitiba; Foz do Iguaçu; Prancheia; Londrina; Apucarana; Mal. Candido Rondon; Medianeira; Arapongas; Francisco Beltrão; Paranaguá; Mandaguari; Telêmaco Borba; Paranavaí; Missal
Sudeste	Rio de Janeiro	Rio de Janeiro; Parati; Niterói; Angra dos Reis; São João do Mereti; Resende; Rio Bonito; Duque de Caxias; Nova Iguaçu; Itaboraí; Macaé; Nilópolis; São Gonçalo; Saquarema; Maricá
Sudeste	São Paulo	São Paulo; Campinas; Ribeirão Preto; Botucatu; Presidente Prudente; São José do Rio Preto; Mauá; Osasco; Guarujá; Mogiguaçu; Jacareí; Araçatuba; Preseidente Vencelau; Bauru; Fernandópolis

Variável Dependente

- Casos de Malária

Variáveis Explanatórias

- População Total

- Renda per Capita
- Gastos com Saúde e Saneamento Básico
- Coeficiente de Gini
- Índice de Desenvolvimento Humano – Educação
- Taxa de Analfabetismo
- Numero de Médicos por 1.000 habitantes
- Gastos com Habitação e Urbanização
- Transferência de Recursos do Governo Federal
- Transferência de Recursos do Governo Estadual

#### Definição das Variáveis

Casos de Malária: número de casos de malária registrados anualmente, pelo SUS, em cada município

População Total: levantamento censitário da população, realizado pelo IBGE.

Renda per Capita: razão entre o somatório da renda de todos os membros da família e o número de membros da mesma.

Gastos com Saúde e Saneamento Básico: recursos municipais aplicados na área de habitação.

Índice de Gini: mede o grau de desigualdade existente na distribuição de indivíduos segundo a renda domiciliar per Capita.

Índice de Desenvolvimento Humano – Educação: índice obtido a partir da taxa de alfabetização e da taxa bruta de frequência à escola.

Taxa de Analfabetismo: percentual de pessoas de 25 ou mais anos de idade analfabetos.

Numero de Médicos por 1.000 habitantes: número de médicos residentes no município por 1000 habitantes.

Gastos com Habitação e Urbanização: despesa por função de habitação e urbanismo, levantadas pelo IBGE.

Transferência de Recursos do Governo Federal: recursos federais aplicados à saúde, transferidos à cada município.

Transferência de Recursos do Governo Estadual: recursos estaduais aplicados à saúde, transferidos à cada município.

Índice da densidade Kernel foi utilizado no final deste estudo para analisar a normalidade dos resíduos. O objetivo de uma regressão não paramétrica é encontrar de uma relação entre as variáveis sem impor nenhum formato. As técnicas utilizadas neste método geralmente se baseiam na estimação da densidade Kernel, onde a função de densidade pode ser representada por :

$$\kappa_h(x) = \frac{1}{h} \cdot k\left(\frac{x}{h}\right)$$

Sendo  $h$  o parâmetro que ajusta o grau de suavidade da função, denominado de janela. O estimador dessa função é dado por  $\tau_h(x) = \frac{1}{n} \sum_{i=1}^n k_h(x - x_i)$  que representa a média das funções Kernel nas observações  $X_i$ . Como a Kernel é uma função de densidade, então o seu estimador também o é, de modo que  $\int \kappa(x) dx \Rightarrow \int \hat{\tau}_h(x) dx = 1$ .

Por outro lado, a função de densidade Kernel é estimada baseando-se em dois parâmetros, na janela ( $h$ ) e na função Kernel ( $K$ ) escolhida. Se a Janela  $h$  e uma função Kernel são corretamente especificadas então a correspondente função de densidade é única para o conjunto de dados.

A análise dos resíduos executada através dessa função, no presente trabalho, mostra também a normalidade desses dados.

Com base em informações fornecidas pelo Instituto Brasileiro de Geografia e Estatística (IBGE) e pelo Instituto de Pesquisas Econômicas Aplicadas (IPEA), para cada um dos 90 (noventa) municípios selecionados, fez-se o levantamento das 10 (dez) variáveis explanatórias, anteriormente citadas. Para a variável explicada (numero de casos de malária) foi executada uma tabulação dos casos registrados para cada Município, durante os 12 (doze)

meses do ano 2004, abrangendo os principais tipos de malária que assolam a região (Vivax, Malariae e Falciparum).

## 6. RESULTADOS

No modelo nacional, considerando a matriz completa (dados dos 90 Municípios), verificou-se que as variáveis mais importantes para explicar os casos de malária no Brasil são: quantidade de Médicos, o coeficiente de Gini, as transferências de Recursos Federais para a saúde e a Taxa de Analfabetismo. O modelo de regressão múltipla obtido, após análise da tabela 2, encontra-se discriminado abaixo.

Tabela 2. Distribuição das variáveis explanatórias do modelo nacional para os casos de malária no Brasil, 2004

Variáveis Explanatórias	B	P – valor
Médicos	- 1,437	0,042
Gini	20,502	0,000
Tran_Fed	1,470	0,002
Tx. Analf.	-0,393	0,00

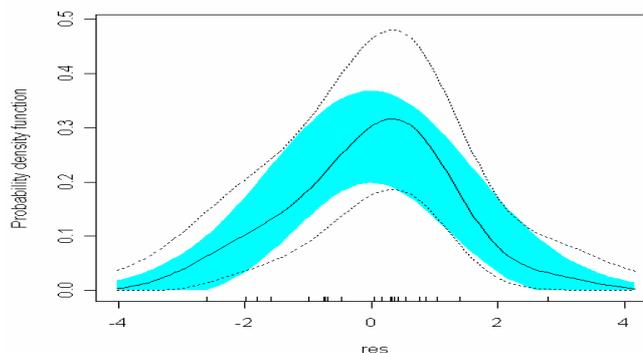
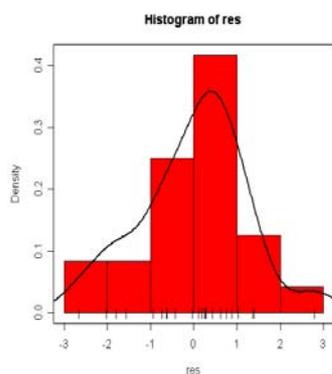
IPEA – Instituto de Pesquisas Econômicas Aplicadas

$$\text{Casos} = -1.437 \text{ Médicos} + 20,502 \text{ Gini} + 1,470 \text{ Tran\_Fed} - 0,393 \text{ Tx. Analf.}$$

$$\text{Com } R^2 = 90,6\%$$

Análise dos Resíduos

Intervalo de Confiança



O modelo para a região Norte, considerando o número de casos de malária como variável dependente, apresentou Gini e Habitação/Urbanização como significativas para explicar a variação no número de incidência de malária:  $b_1$  ( $p < 0,0001$ ) e  $b_3$  ( $p = 0,055$ )

Tabela 3. Distribuição das variáveis explanatórias dos modelos regionais nos casos de malária no Brasil, 2004

Variáveis Explanatórias	Norte		Sudeste		Sul	
	B	P	B	P	B	P
Gini	5,547	0,000	0,107	0,816	1,030	0,031
Hab_Urb	4,702 E - 08	0,055	3,602 E - 09	0,000	0,000	0,011
Médicos	- 0,411	0,378	0,284	0,002	- 4,999	0,714
Tran_Fed	6,916 E - 08	0,797	-6,952 E - 08	0,480	- 8,193 E - 07	0,087
Tran_Est	9,586 E - 07	0,251	-1,651 E - 07	0,611	- 5,076 E - 08	0,895

O modelo da malária para região Norte, têm por equação:

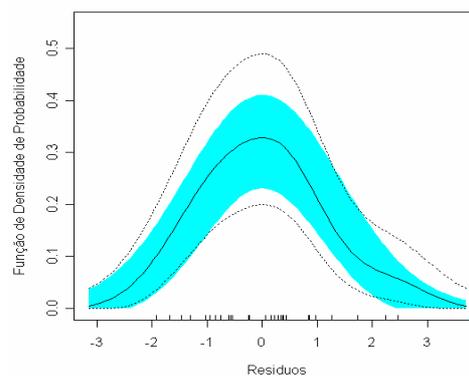
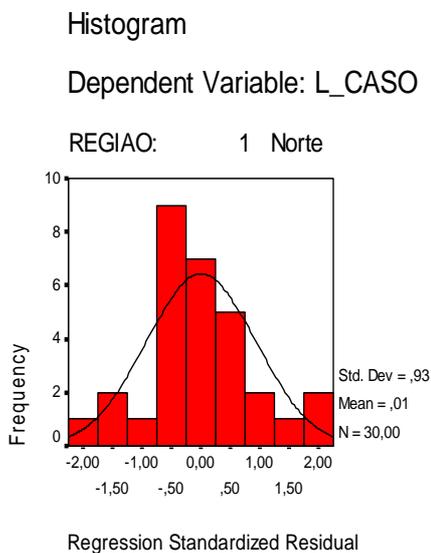
$$\text{Casos} = b_1 \text{ Gini} + b_3 \text{ Hab\_Urb}$$

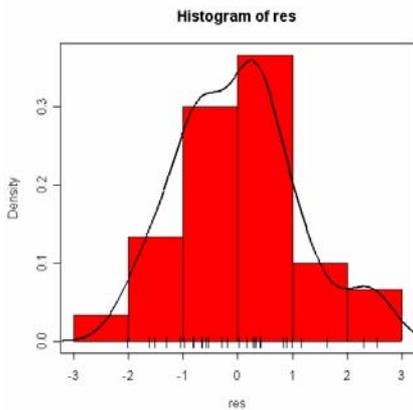
$$\text{Casos} = 5,547 \text{ Gini} + 4,702 \times 10^{-8} \text{ Hab\_Urb}$$

Com  $R^2 = 92,7 \%$

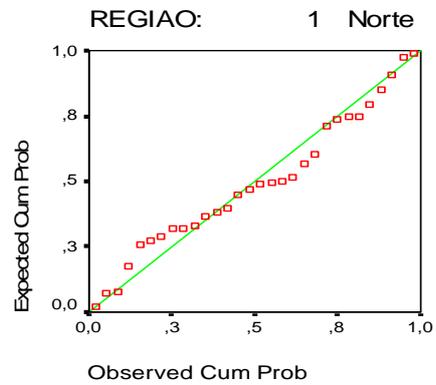
Análise dos Resíduos

Intervalo de Confiança





Normal P-P Plot of Regression  
Dependent Variable: L\_CASO



Para a região Sudeste, as variáveis significativas foram: médicos e Habitação/Urbanização, ( $p = 0,002$ ) e ( $p < 0,0001$ ), respectivamente.

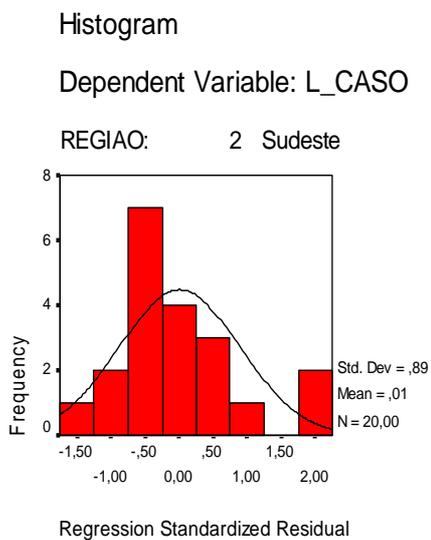
Então:

$$\text{Casos} = b_2 \text{ Médicos} + b_3 \text{ Hab\_Urb}$$

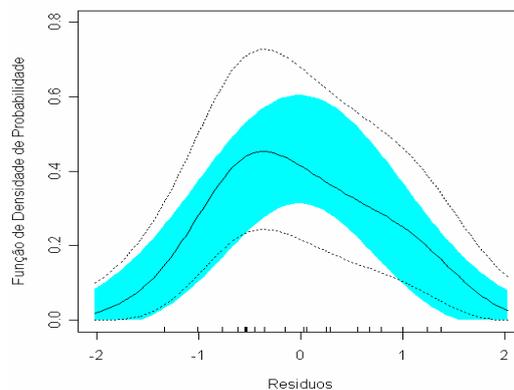
$$\text{Casos} = 0,284 \text{ Medicos} + 3,602 \times 10^{-9} \text{ Hab\_Urb}$$

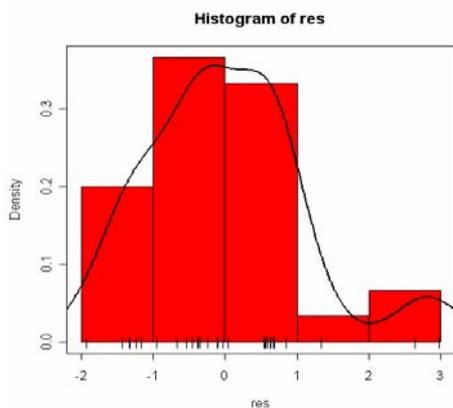
Com  $R^2 = 86,90\%$ .

### Análise dos Resíduos



### Intervalo de Confiança

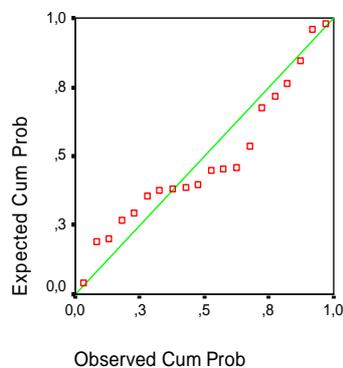




Normal P-P Plot of Regression

Dependent Variable: L\_CASO

REGIAO: 2 Sudeste



Para a região Sul, as variáveis explicativas do Modelo são: Gini e Habitação/Urbanização, ( $p = 0,031$  e ( $p = 0,011$ ), respectivamente

Então:

$$\text{Casos} = b_1 \text{ Gini} + b_3 \text{ Hab\_Urb}$$

$$\text{Casos} = 1,030 \text{ Gini} + 1,567 \times 10^{-8} \text{ Hab\_Urb}$$

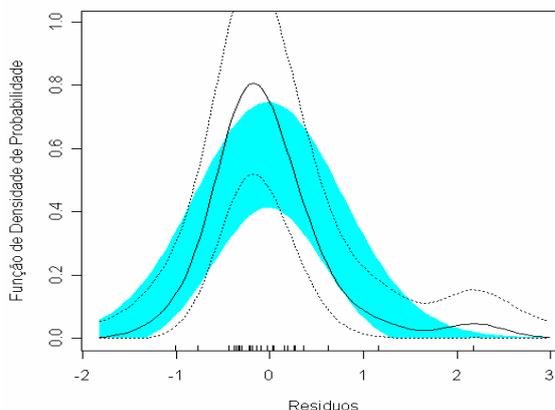
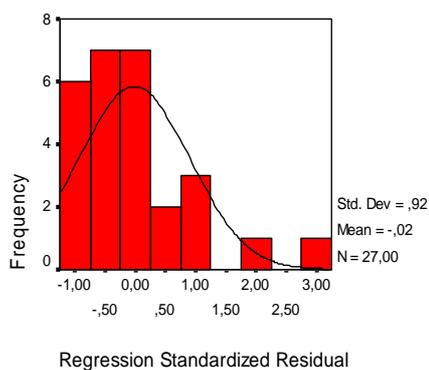
$$\text{têm } R^2 = 59,2\%$$

Nos três modelos a variável Habitação/Urbanização foi considerada significativa para explicar o número de malária em cada região. A diferença entre os mesmos fica por conta por conta da variável Gini (incluída no modelo I e III) e Médicos (incluída no modelo II).

Histogram

Dependent Variable: L\_CASO

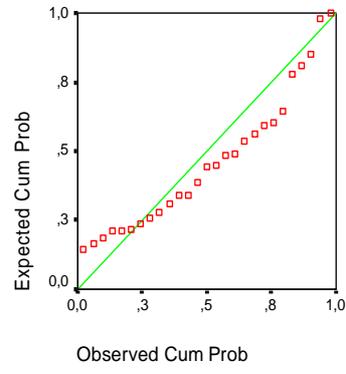
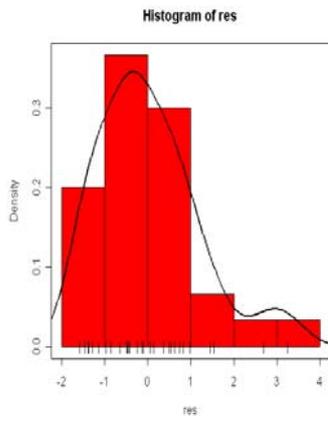
REGIAO: 3 Sul



Normal P-P Plot of Regression

Dependent Variable: L\_CASO

REGIAO: 3 Sul



O modelo da região Norte explica 92,7% dos casos de malária, o da região Sudeste 86,90% e da região Sul 59,2%.

No modelo nacional, os sinais negativos para os coeficiente das variáveis Médicos e Taxa de Analfabetismo, representam uma redução no número de casos da doença. Por outro lado, quanto maior o Gini e os Recursos Federais, verifica-se maior número de casos de malária.

Este fato pode ser explicado pela ineficiência na aplicação de verbas pelos setores públicos.

O modelo da região Norte apresentou o maior coeficiente de determinação ( $R^2 = 92,7\%$ ), e também as variáveis significativas Gini e Habitação/Urbanização o que era esperado, por tratar-se de região subdesenvolvida, com habitação e saneamento precários ou inexistentes, o que provavelmente favorece a ação dos vetores de transmissão da malária.

As variáveis significativas no modelo nacional foram Gini, Taxa de Analfabetismo, Médicos e Transferência de Recursos do Governo Federal.

Nos três modelos regionais, a variável Habitação/Urbanização foi considerada significativa para explicar o número de casos de malária em cada região. A diferença entre os modelos fica por conta apenas da variável Gini (incluída nos modelos Norte e Sul) e Médicos (incluída apenas no modelo Sudeste).

Os resíduos, para as regiões Norte (média de 0,001 e desvio padrão de 0,93) e Sudeste (média de 0,01 e desvio padrão de 0,89) ajustam-se aproximadamente à distribuição normal.

Na região Sudeste, a variável Gini não foi significativa, por tratar-se da região mais desenvolvida do país, apresentando, historicamente, uma baixa concentração de renda. Desta forma, as variáveis significativas selecionadas foram Médicos e Habitação/Urbanização.

Para a região Sul (média de 0,02 e desvio padrão de 0,92) a distribuição de resíduos apresenta forte assimetria à esquerda, implicando no baixo poder explicativo do modelo. As variáveis significativas para esta região foram Gini e Habitação/Urbanização.

O Governo deveria adotar uma política que possibilitasse melhor distribuição de renda no país, principalmente na Região Norte, onde a desigualdade é mais acentuada, exigindo dos gestores públicos mais eficiência na aplicação dos Recursos Federais transferidos.

## REFERÊNCIAS BIBLIOGRÁFICAS

---

- AITKIN, M.; LAIRD, N.; FRANCIS, B. A reanalysis of the Stanford heart transplant data (with discussion). **Journal of the American Statistical Association**, v. 78, p. 264–292, 1983.
- AGUIAR. Consultoria em informática. Duque de Caxias, 2004.  
Disponível em: <http://www.aguiarsoftware.com.br>.
- ANTHEUS. Análise e reconhecimento em impressões digitais. Curitiba, 2002.  
Disponível em: <http://www.arid.com.br/Tecnologia/Arid1.htm>.
- APPES –Associação dos Papiloscopistas de Espírito Santo. Espírito Santo, 2004.  
Disponível em: <http://www.appes.com.br>.
- APPOL–Associação dos Papiloscopistas do Rio de Janeiro. Rio de Janeiro, 2004.  
Disponível em: <http://www.appol.com.br>.
- ARBOUS, A.G.; KERRICH, J.E., Accident statistics and concept of accident proneness. **Biometrics**, v. 7, p. 340–432, 1951.
- BAHLER, C.; HILL, R.R., Jr.; BYERS, R.A.. Comparasion of logistic and Weibull functions: the effect of temperature on cumulative germination of alfafa. **Crop Science**, v. 29, p. 142–146, 1989.
- BARROS, V.R.; ESTEVAN, E.A., On the evaluation of wind power from short wind records. **Journal of Climate and Applied Meteorology**, v. 22, p. 1116–1123, 1983.
- BENNETT, B.M.; BIRCH, B. Sampling inspection tables for comparasion of two groups using the negative binomial model. **Trabajos de Estadística**, v. 15, p. 1–12, 1964.
- BERRY, G. Design of carcinogenesis experiments using the Weibull distribution. **Biometrika**, v. 62, p. 321–328, 1975.
- BHATTACHARYYA, G.K.; FRIES, A. Fatigue Failure Models-Birnbaum-Saunders vs. Inverse Gaussian. **IEEE Transactions in Reliability**, v. 31, p. 439 –440, 1982.
- BINNS, M.R. Behavioral dynamics and the negative binomial distribution. **Oikos**, v. 47, p. 315–318, 1986.
- BLACKMAN, G.E. A study by statistical methods of the distribution of species in a grassland association. *Annals of Botany (New Series)*, **49**, 749–778, 1935.
- BOES, D.C. Regional flood quantile estimation for a Weibull model. **Water Research**, v. 25, p. 979–990, 1986.
- BOSWELL, M.T.; ORD, J.K.; PATIL, G.P. Chance mechanisms underlying univariate Distributions. *Statistical Ecology, 4: Statistical Distributions in Ecological Work*, J.K. Ord, G.P. Patil, and C.Taillie(editors), 1–156. Fairland, MD: International Co-operative Publishing House, 1979.

BRESSAN, N.M. *Biometria*, 2002. Disponível em: [http://www.ucs.br/ccet/demc/vj\\_brussam/inst/biometria.pdf](http://www.ucs.br/ccet/demc/vj_brussam/inst/biometria.pdf).

BURTON, P.W.; MAKROPOULOS, K.C. Seismic risk of circum-Pacific earthquakes: **II**. Extreme values using Gumbel's third distribution and the relationship with strain energy release. **Pure and Applied Geophysics**, v. 123, p. 849–866, 1985.

CARLIN, J.; HASLETT, J. The probability of wind power from a dispersed array of wind turbine generators. **Journal of Climate and Applied Meteorology**, v. 21, p. 303–313, 1982.

CASTRO NETO, P.; SILVEIRA, J.V. Precipitação provável em Lavras, região sul de Minas Gerais, baseada na função de distribuição de probabilidade gama. **Ciência e Prática, Lavras**, v.5, n.2, 144–151, 1981a.

CHEN, W.C.; HILL, B.H.; GREENHOUSE, J.B.; FAYOS, J.V. Bayesian analysis of survival curves for cancer patients following treatment. in *Bayesian Statistics*, 2, J.M. Bernardo, M.H. De Groot, D.V. Lindley, and A.F.M. Smith (editors), 299–328, Amsterdam: North-Holland, 1985.

CHHIKARA, R.S.; FOLKS, J.L. The inverse Gaussian distribution as a lifetime model. **Technometrics**, v. 19, p. 461–468, 1977.

CLIFF, D.; ORD, J. K., *Spatial Processes: Models and Applications*, London: Pion, 1981.

COCHRAN, W.G. Some methods for strengthening the common  $\chi^2$  tests. **Biometrics**, v. 10, p. 417–451, 1954.

CONRANDSEN, K.; NIELSEN, L.B.; PRAHM, L.P. Review of Weibull statistics for estimation of wind speed distributions. **Journal of Climate and Applied Meteorology**, v. 23, p. 1173–1183, 1984.

CORDEIRO, G.M.; NETO, E.A.A.L. *Modelos Paramétricos*. XVI SINAPE, ABE, 1974.

COSTA, S.M. **Classificação e reconhecimento de impressões digitais**. 2000. 123f. Dissertação, (Mestrado em Engenharia Elétrica), Escola Politécnica da Universidade de São Paulo, São Paulo. Disponível em: <http://sim.lme.usp.br/publicações/exames/pdf/QualiSi.pdf>.

COX, D.R.; SNELL, J.E. *Analysis of Binary Data* (second edition). London: Chapman and Hall, 1989.

DAS, S.C. The fitting of truncated type III curves to daily rainfall data. **Australian Journal of Physics**, v. 7, p. 298–304, 1955.

DENNIS, B.; PATIL, G.P. The gamma distribution and the weighted multimodal gamma distributions as models of population abundance. **Mathematical Biosciences**, v. 68, p. 187–212, 1984.

DIXON, J.C.; SWIFT, R.H. The directional variation of wind speed and Weibull probability parameters. **Atmospheric Environment**, 18, p. 2041–2047, 1984.

DOOB, J.L., Stochastic Processes. New York: Wiley, 1953.

DYER, A.R. An analysis of the relationship of systolic blood pressure, serum cholesterol, and smoking to 14-year mortality in the Chicago Peoples Gas Company Study. Part I: total mortality in exponential – Weibull model. PartII: coronary and cardiovascular –renal mortality in two competing risk models. **Journal of Chronic Diseases**, v. 28, p. 265–578, 1975.

FISHER, R.A.; CORBET, A.S.; WILLIAMS, C.B., The relation between the number of species and the number of individuals in a random sample of an animal population. **Journal of Animal Ecology**, v. 12, p. 42–58, 1943.

FISHER, R.A.; MATHER, K. A linkage test with nice. *Annals of Eugenics*, London, **7**, 265–280, 1936.

FRIEDMAN, D.G.; JONES, B.E. Estimation of rainfall probabilities. University of Connecticut Agricultural Experiment. **Station Bulletin, Connecticut**, v. 332, p. 1–22, 1957.

FRIENBERG, S.E. Contingency tables. *Encyclopedia of Statistical Sciences*, S. Kotz, N.L.Johnson and C.B. Read (editors), **2**, 161–171, New York: Wiley, 1982.

FRIZZONE, J.A. Análise de cinco modelos para cálculo da distribuição e frequência de precipitação pluviométrica na região de Viçosa. 1979.

FULLER, W.E. Flood flows. *Transactions of the American Society of Civil Engineers*, **77**, 564, 1914.

FURRY, W.H. Passage of high energy electrons through lead. **Physical Review**, 52, 569–581, 1937.

GALATE, R.S. **Estudo de precipitações pluviiais no município de Belém-PA, através da distribuição gama. Piracicaba: ESALQ/USP.**, (Dissertação de mestrado), 1987.

GALTON, F. The geometric mean in vital and social statistics. *Proceedings of the Royal Society of London*, **29**, 365–367, 1879.

GLADE, J.E. **Prognose de volume por classe diamétrica para Eucalyptus Grandis Hill Ex-Maiden. Curitiba: Dissertação de Mestrado: Universidade Federal do Paraná**, 1986.

GORSKI, A.C. Beware of the Weibull euphoria. **IEEE Transactions on Reability**, v. 17, p. 202–203, 1968.

GREENWOOD, M.; YULE, G.U. An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. **Journal of the Royal Statistical Society**, serie A, v. 83, p. 255–279, 1920.

GRIFFITH, A.A. The phenomena of rupture and flow in solids. *Philosophical Transactions of the Royal Society of London, series A*, **221**, 163–198, 1920

GUMBEL, E.J. Les intervalles extremes entre les emissions radioactives. **Journal de Physique et de Radium**, **8**, 446–452, 1937a.

GUMBEL, E.J. La durée extreme de la vie humaine. *Actualités Scientifiques et Industrielles*, Paris: Hermann et Cie, 1937a.

GUMS, R.A. Protótipo de um sistema de identificação de minúcias em impressões digitais utilizando redes neurais artificiais multicamada. Centro de Ciências Exatas e Naturais, Universidade Regional de Blumenau, Blumenau, 2002.

HAIGHT, F.A. *Handbook of the Poisson Distribution*. New York: Wiley, 1967.

HALL, M.J. Problems of handling messy field data for engineering decision making: More on flood frequency analysis. **The Mathematical Scientist**, v.17, p. 78–88, 1992.

HARROP-WILLIAMS, K. Random nature of soil porosity and related properties. **Journal of Engineering Mechanics**, v. 115, p. 1129–1133, 1989.

HENERY, R.J. An extreme-value model for predicting the results of horse races. **Applied Statistics**, v. 33, p. 1125–1133, 1984.

HERMANSON, R.E.; JOHNSON, H.P. Generalized flood-frequency relationships. **Iowa State Journal of Science**, v. 41, p. 247–268, 1967.

HOMAN, S.M.; PARDO, E.S.; TRUDEAU, M. An application of Weibull modeling and discriminant analysis of alcoholism relapse among veterans. *ASA Proceedings of Social Statistics Seccion*, 222–224, 1987.

HONG, L. **Automatic personal identification using fingerprints**. (Tese de Doutorado em filosofia). Departamento de ciências da computação, Michigan State University, 1998. Disponível em: <http://www.cse.msu.edu/publications/tech/TR/MSUCPS-98-24ps.gz>.

HOPKE, P.K.; PAATERO, P. Extreme value estimation applied to aerosol size distributions and related environmental problems. Paper presented at the Conference on Extreme Value Theory and its Applications, May 2-7, 1993, National Institute of Standards, Gaithersburg, MD, 1993.

JANARDAN, K.G.; PADMANABHAN, G. Double bounded beta distribution for hydrologic variables. *Proc. 17th Annual Pittsburgh Conference (part 3)*, **17**, 1107–1111, 1986.

JENKINSON, F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. **Quarterly Journal of the Royal Meteorological Society**, v. 81, p. 158–171, 1955.

JUNIOR, T.; DA SILVA, G. *A Papiloscopia nos locais do crime*. São Paulo: Ícone, 159p., 1991.

KANAROGLOU, P.; LIAW, K.L.; PAPAGEOGIOU, Y.Y. An analysis of migratory systems. II: Operational framework, *Environment and Planning, series A*, **18**, 1039–1060, 1986.

KAPLAN, N.; RISCO, K. A method for estimating rates of nucleotide substitution using DNA sequence data. *Theoretical Population Biology*, v. 21, p. 318–328, 1982.

KENDALL, D.G., On some modes of population growth leading to R.A.Fisher's logarithmic series distribution. *Biometrika*, v. 35, p. 6–15, 1948.

KENDALL, D.G. Stochastic processes and population growth. *Journal of the Royal Statistical Society, Series B*, v. 11, p. 230–282, 1949.

KING, J.R. Summary of extreme-value theory and its relation to reliability analysis. *Proceedings of the 12th Annual Conference of the American Society for Quality Control*, **13**, 163–167, 1959.

KEHDY, C. *Elementos de Criminalística*. 3.ed. São Paulo: Sugestões literárias, 1968.

KEMP, C.D. Accident proneness and discrete distribution theory, *Random Counts in Scientific Work. 2: Random Counts in Biomedical and Social Sciences*, G.P. Patil (editor), 41–65. University Park: Pennsylvania State University Press, 1970.

KOPPES, L.J.; GROVER, N.B. Relationship between size of parent at cell division and relative size of progeny in *Escherichia Coli*. *Archives of Microbiology*, v. 157, p. 402–405, 1992.

KOTZ, S.; NEUMANN, J. On distribution of precipitation amounts for Geophysical Research. **68**, 3635–3641. Addendum, **69**, 800–801, 1963.

MARGOLIN, B.H.; RESNICK, M.A.; RIMPO, J.H.; ARCHER, P.; GALLOWAY, S.M.; BLOOM, A.D.; ZEIGER, E. Statistical analysis for in vitro cytogenetics assays using Chinese hamster ovary cells. *Environmental Mutagenesis*, v. 8, p. 183–204, 1986.

MAYER, P.L. *Probabilidade Aplicações à Estatística*. Departamento de Matemática. Washington State University, 1975.

McNALLY, R.J. Maximum likelihood estimation of the parameters of the prior distributions of three variables that strongly influence reproductive performance in cows. *Biometrics*, v. 46, p. 501–514, 1990.

NATHAN, R.J.; McMAHON, T.A. Practical aspects of low-flow frequency analysis. *Water Resources Research*, v. 26, p. 2135–2141, 1990.

OKUBO, T.; NARITA, N. On the distribution of extreme winds expected in Japan. *National Bureau of Standards Special Publication*, 560–1, 12pp., 1980.

OLDHAM, D.D. On estimating the arithmetic means of lognormally distributed populations. *Biometrics*, v. 21, p. 235–239, 19??.

PADGETT, W.J.; TSAI, S.K. Prediction intervals for future observations from the inverse Gaussian distribution. *IEEE Transactions on Reliability*, **35**, 406-408, 1986.

PANKANTI, S. Biometrics: promising frontiers for emerging identification market. Howthorne, 2000.

Disponível em: <http://www.cse.msu.edu/publications/tech/TR/MSU-CSE-00-2.os.gz>.

PERRY, J.N. Negative binomial model for mosquitoes. *Biometrics*, v. 40, p. 863, 1984.

PETERSON, C.G. The yearly immigration of young plaice into the Linfjord from the German Sea. *Danish Biological Station Report*, v. 6, p. 5-48, 1986.

POTTER, W.D. Normalcy tests of precipitation and frequency studies of runoff on small watersheds. U.S. Department of Agriculture Technical Bulletin, no. 985, Washington, DC: GPO, 1949.

PROGLOBO. Sensores em Biometria. São Paulo, 1996.

Disponível em: <http://web01.proglobo.pt:8080/backo/html/soluçõesGA.htm>.

RAJAN, K. Extreme value theory and its applications in microstructural science. Paper present at the Conference on Extreme Value Theory and its Applications. May 2-7, 1993, National Institute of Standards, Gaithersburg, MD, 1993.

RAMTZ, S.F.; RIGGS, H.C. Magnitude and frequency of floods in the Columbia River Basin. U.S. Geological Survey, Water Supply Paper, **1080**, 317-476, 1949.

RAO, B.R.; MAZUMDAR, S.; WALLER, J.H.; LI, C.C. Correlation between the numbers of two types of children in a family. *Biometrics*, v. 29, p. 271-279, 1973.

RAWLINGS, J.O.; CURE, W.W. The Weibull functions as a dose-response model to describe ozone effects on crop yields. *Crop Science*, v. 25, p. 807-814, 1985.

RINK, G.; DELL, T.R.; SWITZER, G. BONNER, F.T. Use of the [three-parameter] Weibull function to quantify sweetgum germinating data. *Silva Genetica*, v. 28, p. 9-12, 1979.

ROSS, R. Ciclo de vida do parasita da Malária. 1987.

ROWE, J.A. Mosquito light trap catches from ten American Cities. 1940, IOWA, State College *Journal of Science*, v. 16, p. 487-518, 1942.

SALEM, A.B.; MOUNT, T.D. A convenient descriptive model of income distribution: the gamma density. *Econometrica*, v. 42, p. 1115-1127, 1974.

SCHWENKE, J.R. Using the Weibull distribution in modeling pharmacokinetic data. ASA Proceedings of the Biopharmaceutical Section, 104-109, 1987.

SCHUETTE, T.; SALKA, O.; ISRAELSSON, S. The use of Weibull distribution for thunderstorm parameters. *Journal of Climate and Applied Meteorology*, v. 26, p. 457-463, 1987.

- SEAL, H.L. Mortality data and the binomial probability law. **Skandinavisk Aktuarietidskrift**, v. 32, p. 188–216, 1949a.
- SEBER, G.A.F. The Estimation of Animal Abundance(second edition), London: Griffin, 1982b.
- SHEN, H.W., BRYSON, M.C.; OCHOA, I.D. Effect of tail behavior assumptions on flood predictions. *Water Resources Research*, **16**, 361–364, 1980.
- SICHEL, H.S. The estimation of the parameters of a negative binomial distribution with special reference to psychological data. **Psychometrika**, v. 16, p. 107–127, 1951.
- SIMIU, E.; BIETRY, J.; FALLIBLE, J.J. Sampling errors in estimation of extreme winds. **Journal of the Structural Division of the National Bureau of Standards**, v. 104, p. 491–501, 1978.
- SIMIU, E.; FILLIBEN, J.J. Statistical analysis of extreme winds. National Bureau of Standards Technical Note, 868, 52pp., 1975.
- SIMIU, E.; FILLIBEN, J.J. Probability distributions of extreme wind speeds. **Journal of the Structural Division of the National Bureau of Standards**, v.102, p.1861-1877, 1976.
- SOONG, T.T. Modelos Probabilísticos em Engenharia e Ciências. Universidade Estadual de New York, em Buffalo, New York. ICEX-UFMG, 1986.
- STEYN, H.S. On the univarable serie  $F(t) = F(a; b_1, b_2, \dots, b_k; t_1, t_2, \dots, t_k)$  and its applications in probability theory. *Proceedings Koninklike Nederlandse.Akademie van wetenschappen*, series A, **59**, 190–197, 1956.
- TAYLOR, H.M.; KARLIN, S. An Introduction to Stochastic Modeling. Orlando, FL: Academic Press, 1984.
- THOM H.C.S. Frequency of maximum wind speeds. *Proceedings of the American Society of Civil Engineers*, **80**, 104-114, 1954.
- TWEEDIE, M.C.K. Functions of a statistical variate with given means, with special reference to Laplacian Distribution. *Proceedings of the Cambridge Philosophical Society*, **43**, 41–49, 1947.
- VAN DER AUWERA, L.; DE MEYER, F.; MALET, L.M. The use of the Weibull three parameter model for estimating mean wind power densities. **Journal of Applied Meteorology**, v. 19, p. 819–825, 1980.
- VELZ, C.J. Factors influencing self-purification and their relation to pollution abatement. **Sewage Works Journal**, v. 19, p. 629–644, 1947.
- VODA, V.G. Concerning an application of the Weibull distribution to reliability aspects in metrology, *Metrologia Applicata*, **25**, 123–124, 1978.

ZHANG, Y. Plotting positions of annual flood extremes considering extraordinary values. **Water Resources Research**, v. 18, p. 859–864, 1982.

WASAN, M.T. First Passage Time Department of Mathematics. Kingston, Ontario: Queen's University, 1968b.

WHITTEMORE, A.; ALTSCHULER, B. Lung cancer incidence in cigarette smokers: Further analysis of Doll and Hill's data for British Physicians. **Biometrics**, v. 32, p. 805–816, 1976.

WILEY, J.A.; HERSCHOKORU, S.J.; PADIAU, N.S. Heterogeneity in the probability of HIV transmission per sexual contact: The case of male-to-female transmission in penile-vaginal intercourse. **Statistics in Medicine**, v. 8, p. 93–102, 1989.

WILLIAMS, C.B. The number of publications written by biologists. *Annals of Engenics*, **12**, 143–146, 1944.

WILLIAMS, C.B. The logarithmic series and its application to biological problems. **Journal of Ecology**, v. 24, p. 253–272, 1947.

WITHERBY, H.F.; JOURDAIN, F.C.R.; TICHIHURST, N.F.; TUCKER, B.W. The Handbook of British Birds, **5**, London; H.T., and G. Witherby, 1941.

WONG, R.K.W. Weibull distribution, interative likelihood techniques and hidrometeorological data. **Jounal of Applied Meteorology**, v.16, p. 1360–1364, 1977.



Tabela 1. Distribuição do Número de Casos Notificados de Malária X Variáveis Sócio-econômicas por Municípios das Regiões Norte, Sudeste e Sul, Brasil, 2004

	Nº. de Casos	Renda per Capita	Populaçã o Total	Saúde e San.	Gini	IDH-Ed	Tx. Analf.	Nº. Medicosp/mil	Habitação/Uurb.	Transf. Fed	Transf. Est
<b>PARÁ</b>	Y	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
Altamira	130	205	62.285	9.165.028	0.592	0.797	30.2	0.63	8.016.004	35.393	0
Belém	38	313.93	1.272.354	200.851.724	0.651	0.928	6.9	2.1	28.018.810	873.377	0
Marabá	50	188.59	134.373	17.859.298	0.632	0.799	23.7	0.4	7.192.799	833.855	760.028
Santarém	59	139.9	186.297	19.701.070	0.618	0.884	17.2	0.38	5.181.294	3.963.112	0
Itaituba	405	162.6	64.486	3.978.979	0.649	0.797	33.9	0.23	15.828	92.240	0
Jacareacanga	66	140.92	5.670	0	0.644	0.615	0	0	0	0	0
Tucuruí	43	180.99	60.918	7.881.357	0.592	0.867	26	0.83	2.248.312	245.105	0
Anajás	13	89.82	4.613	0	0.556	0.529	64.9	0	0	0	0
Santana	257	121.08	17.326	0	0.618	0.751	34.3	0	0	0	0
Castanhal	14	162.25	121.249	9.440.432	0.573	0.854	17.5	0.27	4.713.655	372.188	0
Dom eliseu	4	145.7	23.801	2.544.074	0.65	0.726	44.6	0	1.263.420	294.750	60.000
Paragominas	5	166.04	58.240	0	0.606	0.766	38	0.34	0	0	0
Redenção	69	200.72	59.613	7.042.532	0.605	0.836	25.5	0	2.286.022	620.692	454.291
Itupiranga	61	85.71	14.754	2.481.718	0.602	0.671	46.8	0.21	1.406.690	669.217	0
Monte Alegre	14	87.87	20.921	68.028	0.59	0.784	24.9	0.32	354.449	1.381.626	0
<b>AMAZONAS</b>											
Boca do acre	69	92.58	26.959	1.734.185	0.623	0.637	53.7	0	858.741	28.818	0
Manaus	667	262.4	1.405.835	68.841.915	0.639	0.909	8.6	0.8	99.996.899	6.563.705	1.940.03
Lábrea	45	66.46	28.956	1.062.929	0.69	0.633	58.5	0	93.419	28.818	3
Apuí	72	130.08	13.864	920.610	0.762	0.778	32.5	0	296.952	613	0
Manacapurú	18	101.23	73.695	0	0.59	0.761	34.3	0	0	0	0
Careiro	38	72.41	27.554	1.218.307	0.66	0.709	51.7	0	1.270.967	28.818	0
Borba	9	60.61	28.619	1.029.544	0.668	0.734	38.9	0	1.157.283	69.791	0
Tefê	27	117.55	64.457	0	0.593	0.75	30.9	0	0	0	0
São Gabriel	36	106.61	29.947	2.658.711	0.806	0.74	29.2	0.27	1.413.561	128.079	338.000
Urucará	55	132.73	18.372	1.633.061	0.766	0.8	18.8	0	1.005.426	53.642	0
Barcelos	7	68.85	24.197	1.407.991	0.655	0.537	41	0	1.083.369	1.057.454	1.137.59
Tabatinga	25	142.08	37.919	2.813.418	0.641	0.78	28.6	0	2.153.573	2.072.889	4
Novo Airão	9	93.27	9.651	807.289	0.584	0.771	33.5	1.81	555.089	90.255	458.431
Iranduba	28	113.85	32.303	2.127.159	0.559	0.766	27.8	0	484.284	272.518	0
Cadajás	7	70.9	17.507	757.799	0.532	0.679	48.8	0	679.546	108.775	0

<b>SÃO PAULO</b>	Y	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
São Paulo	46	610.04	10.434.252	990.610.375	0.618	0.919	4.9	2.7	995.987.427	919.686	2.153.193
Campinas	11	614.86	969.396	188.598.384	0.584	0.925	5	3.43	54.282.636	0	233.418
Ribeirão Preto	10	539.84	504.923	178.576.664	0.564	0.918	4.4	4.68	13.992.434	1.588.622	318.750
Botucatu	2	426.18	108.306	4.700.134	0.539	0.909	5.6	3.74	6.337.118	0	1.421.360
Presidente Prudente	2	482.62	189.186	13.514.842	0.592	0.924	6.2	2.56	7.870.710	162.497	329.927
S. José do Rio Preto	3	512.01	358.523	67.113.072	0.557	0.916	5.4	3.65	10.440.369	3.296.099	419.711
Mauá	2	274.82	363.392	46.289.950	0.49	0.909	6.6	0.05	30.853.650	0	445.102
Osasco	1	390.45	652.593	100.746.145	0.522	0.913	5.8	0.51	3.021.103	544.659	39.170
Guarujá	2	308.41	264.812	33.891.820	0.525	0.885	8.4	0.69	39.854.124	0	209.985
Mogi-Guaçu	1	453.33	124.228	31.535.951	0.519	0.886	7.1	0.69	9.455.674	536.000	576.000
Jacareí	1	353.95	191.291	46.644.843	0.528	0.913	6	0.8	20.987.098	430.620	88.977
Araçatuba	1	503.17	169.254	56.088.438	0.634	0.909	6.3	2.14	12.576.699	5.150.000	21.840
Presidente Vencelau	1	364.66	37.347	3.685.724	0.597	0.893	9.7	0.8	1.911.612	790.287	995.358
Bauru	1	500.27	316.064	43.704.771	0.585	0.908	5.2	1.35	24.287.115	172.107	180.000
Fernandópolis	1	366.5	61.647	4.260.736	0.541	0.891	9.2	1.67	2.241.073	288.000	82.090
<b>RIO DE JANEIRO</b>											
Rio de Janeiro	18	596.65	5.857.904	828.097.433	0.616	0.933	4.4	3.68	385.126.048	450.000	0
Parati	1	312.07	29.544	4.624.791	0.594	0.827	12.8	0.4	1.089.385	0	0
Niterói	4	809.18	459.451	82.892.564	0.587	0.96	3.6	7.27	74.000.493	0	0
Angra dos Reis	1	275.66	119.247	35.583.431	0.553	0.87	8.9	0.72	4.200.714	2.764.114	92.000
S. João do Meretú	1	233.12	449.476	32.713.472	0.475	0.895	5.7	0.2	12.017.744	2.898.377	0
Resende	0	365.45	104.549	21.094.469	0.565	0.918	6.9	1.93	5.826.621	1.233.001	158.931
Rio Bonito	0	276.19	49.691	2.726.191	0.574	0.833	12.8	1.56	2.897.896	1.047.412	60.718
Duque de Caxias	0	226.14	775.456	87.343.897	0.528	0.873	8	0.2	29.444.592	220.000	50.074
Nova Iguaçu	0	237.5	920.599	73.949.921	0.526	0.884	7.2	0.89	29.738.159	7.758.792	39.973.800
Itaboraí	0	202.29	187.479	21.249.420	0.521	0.884	10.8	0.46	3.783.274	0	0
Macaé	0	392.94	132.461	17.637.142	0.562	0.889	7.9	1.48	31.331.151	0	0
Nilópolis	0	298.3	153.712	7.702.620	0.479	0.933	5.7	0.44	2.751.281	2.753.797	0
São Gonçalo	0	268.79	891.119	58.871.177	0.489	0.896	5.8	0.36	21.752.118	0	0
Saquarema	0	266.64	52.461	10.272.548	0.551	0.848	12	0.63	1.482.040	2.615.055	6.316.115
Marica	0	321.41	76.737	5.893.003	0.549	0.881	8.7	0.63	3.966.433	90.000	0
Caxias do Sul	2	490.65	360.419	71.900.922	0.511	0.945	3.7	2.03	112.657	177.059	510.162

<b>PARANÁ</b>	Y	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
Curitiba	7	619.82	1.587.315	107.701.036	0.594	0.946	3.4	2.98	191.897.099	1.708.000	2.154.373
Foz do Iguaçu	12	326.19	258.543	26.119.487	0.582	0.905	7.5	0.77	17.906.123	12.546	58.700
Pranchita	5	254.47	6.260	698.298	0.584	0.855	11.4	1.39	44.409	70.000	103.294
Londrina	1	439.35	447.065	110.208.024	0.577	0.91	7.1	2.16	6.284.351	95.912	1.069.051
Apucarama	1	288.47	107.827	1.888.614	0.513	0.877	9.6	0.82	5.052.465	196.600	10.000
Umuarama	2	313.76	90.690	15.110.164	0.537	0.898	9.8	1.42	3.463.195	390.088	386.517
Mal. Candido											
Rondon	1	341.71	41.007	4.777.760	0.567	0.932	4.3	1.96	5.930.630	35.063	0
Medianeira	1	277.5	37.827	1.912.808	0.513	0.904	7.2	1.08	1.188.248	45.051	34.564
Arapongas	1	304.07	85.428	4.527.118	0.504	0.883	8.4	0.65	6.570.150	677.169	204.382
Franc. Beltrão	1	307.97	67.132	10.442.384	0.577	0.918	8.4	0.85	3.587.493	840.827	122.642
Paranaguá	1	305.36	127.339	8.861.291	0.563	0.897	5.9	0.55	3.571.266	8.492	60.295
Mandaguari	1	238.95	31.395	1.531.968	0.502	0.888	11.8	0.3	160.495	0	0
Telemaco Borba	1	275.6	61.238	3.612.343	0.58	0.865	10.5	0.92	3.611.809	651.000	76.809
Paranavaí	1	312.97	75.750	4.438.090	0.567	0.886	9.8	1.68	1.135.285	330.720	329.757
Missal	1	237.91	10.433	1.659.870	0.549	0.898	7.8	0	936.972	255.000	0
<b>RIO GRANDE DO SUL</b>											
Porto Alegre	4	709.88	1.360.590	690.428.164	0.606	0.951	3.6	5.93	104.207.038	496.871	41.989
Caxias do Sul	2	490.65	360.419	71.900.922	0.511	0.945	3.7	2.03	112.657	177.059	510.162
Passo Fundo	2	405.65	168.458	11.830.539	0.585	0.912	5.6	2.64	6.102.389	15.000	236.165
Frederico											
Westphalen	2	297.09	26.759	2.198.041	0.551	0.937	7.6	1.25	652.048	482.003	287.544
Estrela	1	332.79	27.401	1.577.862	0.469	0.929	3.6	0.72	696.319	198.283	220.206
Maximiliano de											
Almeida	0	192.34	5.651	579.431	0.531	0.883	10.8	0.85	256.593	149.641	25.000
Torres	2	317.74	30.880	2.239.220	0.531	0.911	6.3	1.8	1.780.991	0	342
Ronda Alta	1	245.54	10.051	1.314.708	0.575	0.867	9.4	0.4	195.813	602.930	39.488
Alegria	2	160.89	5.367	930.560	0.563	0.854	12.3	0	90.857	20.000	51.000
Bento Gonçalves	1	466.84	91.486	6.011.847	0.473	0.938	3.9	0.73	2.197.841	275.457	433.468
Lajeado	1	395.34	64.133	4.631.319	0.506	0.923	4.1	0.84	2.235.892	216.630	79.871
Ijuí	1	332.05	78.461	4.752.438	0.572	0.926	5.7	1.28	1.603.367	710.196	327.550
Taquara	1	345.44	52.825	4.285	0.524	0.896	6.8	1.25	158.914	37.644	18.092
Sapucaia do Sul	0	271.38	122.751	6.385.902	0.464	0.9	5.5	0.24	2.868.537	50.213	0
São Leopoldo	0	370.06	193.547	40.757.529	0.551	0.922	4.8	1.28	6.275.356	400.000	967.529

MALÁRIA (Todas as Formas) - Lâminas Positivas, por Unidade Federada, Brasil, 1980 - 2003

UF/Microrregião	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	Total
<b>BRASIL</b>	<b>188.871</b>	<b>187.148</b>	<b>201.828</b>	<b>207.837</b>	<b>279.227</b>	<b>388.482</b>	<b>443.837</b>	<b>588.884</b>	<b>558.535</b>	<b>577.539</b>	<b>588.388</b>	<b>541.807</b>	<b>573.883</b>	<b>493.387</b>	<b>585.135</b>	<b>584.579</b>	<b>644.044</b>	<b>485.026</b>	<b>471.884</b>	<b>837.474</b>	<b>615.267</b>	<b>388.779</b>	<b>348.865</b>	<b>485.017</b>	<b>18.758.771</b>
<b>NORTE</b>	<b>136.727</b>	<b>150.527</b>	<b>184.841</b>	<b>248.723</b>	<b>289.832</b>	<b>323.882</b>	<b>350.742</b>	<b>486.778</b>	<b>475.587</b>	<b>481.261</b>	<b>367.287</b>	<b>388.886</b>	<b>367.085</b>	<b>321.187</b>	<b>438.218</b>	<b>484.117</b>	<b>383.780</b>	<b>286.124</b>	<b>439.743</b>	<b>388.888</b>	<b>522.658</b>	<b>281.884</b>	<b>325.474</b>	<b>289.582</b>	<b>3.765.421</b>
Roraima	59.178	50.589	58.938	88.722	151.140	189.682	191.727	220.969	276.439	244.939	174.330	189.472	132.956	111.290	136.923	134.231	84.054	80.182	77.985	83.298	54.074	57.679	71.224	83.914	2.968.225
Acre	6.180	6.582	7.273	11.860	11.785	10.120	11.959	21.943	21.428	15.785	14.455	13.941	12.613	16.481	26.548	38.063	13.968	6.125	9.843	23.730	21.580	7.774	9.216	12.104	352.542
Amazonas	4.487	6.189	13.142	18.289	9.528	11.198	15.319	15.253	19.382	34.944	26.479	45.949	37.885	55.384	68.267	52.602	70.044	94.382	114.748	107.722	98.028	48.365	78.223	137.532	1.529.760
Roraima	13.028	11.882	15.917	16.891	15.940	15.898	10.927	11.170	14.781	21.019	24.927	25.198	23.984	19.249	24.954	39.574	35.471	25.949	21.263	36.238	35.074	18.028	9.036	11.781	496.878
Pará	38.023	52.450	71.168	116.898	128.930	113.649	135.725	119.120	125.744	127.941	139.736	102.235	129.291	120.041	155.432	185.983	148.334	144.012	179.572	248.223	279.254	196.367	148.888	114.987	3.248.864
Amazônia	1.731	3.130	5.138	4.255	7.270	7.481	8.219	7.481	9.575	11.158	10.677	9.083	8.989	8.819	12.254	17.034	19.058	24.989	22.589	28.846	35.278	26.487	18.257	18.843	315.723
Tocantins	3.680	5.726	8.265	8.710	8.219	6.929	6.937	5.985	6.259	5.679	4.673	4.199	3.417	2.556	2.029	3.718	2.278	1.633	2.758	2.831	1.640	1.244	1.130	1.201	88.724
<b>NORDESTE</b>	<b>21.057</b>	<b>24.534</b>	<b>14.468</b>	<b>24.785</b>	<b>39.528</b>	<b>42.885</b>	<b>41.287</b>	<b>43.287</b>	<b>44.143</b>	<b>42.824</b>	<b>38.844</b>	<b>36.881</b>	<b>18.881</b>	<b>18.122</b>	<b>38.454</b>	<b>33.216</b>	<b>34.588</b>	<b>21.208</b>	<b>28.883</b>	<b>55.382</b>	<b>79.288</b>	<b>34.818</b>	<b>18.818</b>	<b>11.281</b>	<b>868.424</b>
Maranhão	18.141	22.828	17.899	21.873	24.488	37.989	34.689	36.627	39.253	37.725	34.965	18.781	18.389	18.222	29.588	32.919	31.125	20.981	28.271	64.830	78.918	39.587	18.800	10.928	758.985
Pernambuco	797	1.012	690	1.173	2.385	3.080	4.240	4.147	2.983	3.180	2.438	652	685	538	449	268	134	129	132	159	277	133	64	145	28.828
Ceará	288	329	306	376	612	959	1.020	1.183	915	845	774	281	229	170	130	149	102	79	68	90	135	64	464	81	8.859
Rio Grande do Norte	58	53	99	187	132	180	187	134	145	140	87	122	30	28	28	14	23	33	25	122	34	22	13	12	1.581
Piauí	29	44	48	40	80	131	149	151	153	115	70	49	38	21	29	21	14	14	15	30	2	3	3	5	1.284
Pernambuco	198	85	173	121	123	152	272	370	184	120	132	88	80	44	52	31	65	41	27	46	38	19	21	41	2.882
Alagoas	18	10	13	32	38	29	44	35	41	38	19	17	19	13	5	11	9	10	11	12	8	12	10	10	448
Sergipe	20	9	7	11	22	10	25	19	15	20	12	0	17	5	3	6	8	9	1	6	7	5	8	5	268
Bahia	582	144	214	900	575	624	580	539	585	649	378	527	183	111	168	155	121	55	47	53	70	88	42	84	7.582
<b>SUDESTE</b>	<b>1.787</b>	<b>1.734</b>	<b>1.828</b>	<b>2.864</b>	<b>3.218</b>	<b>3.778</b>	<b>4.547</b>	<b>5.038</b>	<b>4.872</b>	<b>4.152</b>	<b>3.883</b>	<b>1.858</b>	<b>1.880</b>	<b>1.173</b>	<b>1.258</b>	<b>1.127</b>	<b>478</b>	<b>798</b>	<b>684</b>	<b>698</b>	<b>647</b>	<b>584</b>	<b>568</b>	<b>433</b>	<b>48.888</b>
Minas Gerais	287	383	408	668	689	1.061	1.142	979	980	994	647	435	347	225	282	284	178	259	158	179	187	167	172	114	11.588
Espírito Santo	149	117	141	134	245	289	524	749	484	780	315	288	258	188	243	187	173	128	182	97	85	74	8	0	5.784
Rio de Janeiro	89	182	108	187	170	216	189	241	178	183	132	89	127	70	71	54	61	50	68	51	72	78	111	71	2.743
São Paulo	1.282	1.212	1.275	1.887	1.884	2.292	2.883	3.067	2.552	2.215	1.789	1.189	980	712	854	894	687	323	363	282	289	289	225	242	28.888
<b>SUL</b>	<b>1.588</b>	<b>1.482</b>	<b>883</b>	<b>1.214</b>	<b>1.818</b>	<b>1.888</b>	<b>2.782</b>	<b>3.588</b>	<b>3.523</b>	<b>3.838</b>	<b>3.889</b>	<b>2.218</b>	<b>1.888</b>	<b>853</b>	<b>965</b>	<b>683</b>	<b>588</b>	<b>458</b>	<b>384</b>	<b>377</b>	<b>388</b>	<b>387</b>	<b>388</b>	<b>325</b>	<b>38.888</b>
Pernambuco	1.148	1.043	925	1.111	1.473	1.739	2.429	3.182	3.088	4.111	3.016	1.810	1.437	894	539	571	455	379	252	324	254	148	225	188	18.888
Santa Catarina	438	280	78	136	89	81	114	180	189	252	253	181	149	87	57	75	59	34	44	27	50	42	57	58	2.885
Rio Grande do Sul	0	189	84	97	148	177	241	288	329	676	583	245	234	82	69	37	38	37	28	26	38	19	13	0	3.878
<b>CENTRO-OESTE</b>	<b>18.728</b>	<b>18.572</b>	<b>18.853</b>	<b>18.241</b>	<b>18.884</b>	<b>18.881</b>	<b>24.188</b>	<b>27.225</b>	<b>22.229</b>	<b>24.184</b>	<b>147.482</b>	<b>148.188</b>	<b>202.227</b>	<b>121.822</b>	<b>87.454</b>	<b>84.127</b>	<b>38.228</b>	<b>8.482</b>	<b>12.528</b>	<b>11.278</b>	<b>12.525</b>	<b>7.178</b>	<b>7.378</b>	<b>4.828</b>	<b>1.188.428</b>
Mato Grosso do Sul	1.049	890	696	716	1.082	1.385	1.649	1.731	1.433	1.277	975	719	689	381	408	272	111	88	84	83	239	233	98	88	18.272
Mato Grosso	15.287	15.654	15.910	15.284	18.989	18.892	19.147	21.829	27.759	28.711	143.983	145.983	200.746	129.711	88.209	84.889	37.628	15.978	11.988	10.960	11.787	8.932	7.885	4.981	1.095.888
Goiás	2.480	1.811	1.965	2.882	2.581	2.188	2.730	2.982	2.582	3.581	2.159	1.581	1.067	740	884	582	384	243	263	280	180	52	137	17	33.784
Distrito Federal	0	289	184	379	489	579	784	837	885	475	283	225	220	178	174	117	97	189	78	80	58	58	0	0	8.880

FONTE: MS/VS/S, SEC e SINAN a partir de 1999.

[3] O ESTADO DE TOCANTINS FOI CRIADO EM 1988. O SIGLA DE BAHIA EM 1988 FOI BA.

\* Dados preliminares.