CENTRO DE ESTATÍSTICA APLICADA – CEA – USP

RELATÓRIO DE CONSULTA

TÍTULO DO PROJETO: **“**Padrões de ectoparasitismo em aves de cerrado e mata de galeria do Distrito Federal - Brasil**”**

PESQUISADORA: Mieko Ferreira Kanegae

ORIENTADOR: Dr. Miguel Ângelo Marini

INSTITUIÇÃO: Instituto de Biociências – USP

FINALIDADE: Publicação

PARTICIPANTES DA ENTREVISTA: Mieko Ferreira Kanegae
Carlos Alberto de Bragança pereira
Claudia Monteiro Peixoto
Aurora Kyoko Nakati
Marco Aurélio Hirata
César Henrique Torres

DATA: 23/08/2005

FINALIDADE DA CONSULTA: Assessoria na análise dos dados

RELATÓRIO ELABORADO POR: Aurora Kyoko Nakati
Marco Aurélio Hirata

1 – Introdução

As aves são hospedeiras de inúmeros tipos de parasitas que apresentam diversos graus de patogenicidade. Os ectoparasitas são parasitas, externos ao corpo do hospedeiro, que afetam as aves de forma direta através da ingestão de tecidos, causando anemia, danos à pele e inoculando toxinas.

Apesar das parasitoses serem responsáveis por extinções de espécies silvestres, as suas presenças são apenas detectadas ou observadas após causarem danos ao ambiente. Apenas recentemente estão sendo reconhecidas os danos provocados pelos parasitas e patógenos nas populações silvestres. Entre as principais causas pode-se destacar o desmatamento e a fragmentação, que provocam um aumento do contato com espécies de hábitats adjacentes, como animais domésticos. Além disso, a reintrodução de espécies pode proporcionar a entrada de novos patógenos e parasitos que podem produzir conseqüências ainda mais desastrosas (Cunningham 1995).

A ocorrência de ectoparasitas nas aves pode variar em função de diversos fatores ecológicos tais como, a história filogenética dos diferentes táxons de aves, o comportamento social, o local de forrageamento e o tipo de ninho.

Os ectoparasitas considerados neste estudo são os ácaros de plumícolas, carrapatos e malófagos.

Os ácaros de plumícolas são ectoparasitas obrigatórios que vivem exclusivamente nas aves, podendo ocorrer em diversas partes da plumagem, principalmente nas pena de vôo e nas coberteiras das asas.

Os carrapatos são quase onipresentes nos vertebrados terrestres, sendo de grande importância como reserva e vetores de organismos patogênicos (Flechtmann 1990). O seu ciclo de vida engloba três instares: larva, ninfa e adulto, sendo que cada muda é precedida de uma refeição com sangue.

O projeto tem como objetivo: estimar a incidência (ou prevalência) de ectoparasitismo por carrapatos e ácaros plumícolas em aves de cerrado e mata de galeria.

O objetivo deste relatório é apresentar, baseado nas amostras coletadas, as estimativas das incidências (ou prevalências) desses organismos na população das aves na região de observação; Fazenda Água Limpa.

2 - Descrição do Experimento

Os dados foram coletados de janeiro a dezembro de 2002, na Fazenda Água Limpa, localizada a 20 km a sudeste de Brasília.

O estudo foi desenvolvido em dois ambientes: mata de galeria e cerrado. Cada um dos ambientes foi amostrado pelo menos quatro dias por mês. A captura das aves foi realizada com, cerca de, 15 redes ornitológicas de 12x2 metros, dispostas em transectos lineares, totalizando 180m de redes em cada local. As redes foram abertas nas primeiras horas do dia, após o nascer do sol e fechadas às 13:00. As aves capturadas foram marcadas com anilhas metálicas, cedidas pelo CEMAVE/IBAMA.

Em todas as aves capturadas mediram-se os números de carrapatos, de ácaros plumícolas e de malófagos. Os carrapatos foram contados na região da cabeça, pescoço e ventre das aves, soprando-se as penas até que a pele pudesse ser visualizada.

Para verificar a presença e quantificar os ácaros plumícolas, foi realizada uma inspeção das penas das asas e cauda contra a luz do sol. As amostras de ácaros obtidas foram retiradas de diferentes locais das penas. Os ectoparasitas foram coletados com uma pinça e conservados em álcool 70%, sendo posteriormente enviados para identificação no Laboratório de Ixodides da IOC/FIOCRUZ (Fundação Oswaldo Cruz).

3 – Descrição das variáveis

Para cada ave capturada, anotava-se a observação de um vetor (c, a, m) cujos elementos assumiram os valores 0 ou 1 de acordo com a ausência ou presença de carrapatos (c), ácaros (a) e malófagos (m).

4 – Situação do Projeto

A tese da pesquisadora já foi defendida, aprovada e seus resultados submetidos a publicação. A crítica dos revisores da revista deveu-se a ausência, no trabalho submetido, de uma análise estatística adequada.

5 – Restrições e Notação

Considerações iniciais devem ser observadas para que o modelo utilizado seja validado. A primeira delas é a de que toda ave do ambiente tenha a mesma chance de ser capturada. A segunda é a de que não exista preferência de espécies no que diz respeito ao ataque dos diversos tipos de parasita. A última, e talvez a mais rigorosa, é a de que a captura das aves foi feita de forma independente. Isto é, o fato de um tipo de ave ser selecionado não aumenta ou diminui a chance de uma outra ave, de mesma espécie, ser selecionada. Em resumo, a coleta é feita segundo um modelo de seleção independente e identicamente distribuída – iid.

A seguir, a notação utilizada é apresentada:

1. As freqüências amostrais são denotadas por y_{cam}. Isto é, o número de aves com vetor de incidência (c,a,m). Por exemplo, y_{011} é o número de unidades amostrais com ausência de carrapatos e presença de ácaro e malófagos. A Tabela 1 descreve todas as freqüências observadas. O vetor das observações será então
   $Y = (y_{000}, y_{001}, y_{010}, y_{100}, y_{011}, y_{101}, y_{110}, y_{111})$.
2. As proporções populacionais de aves em cada uma das 8 possíveis categorias, (c,a,m), são denotadas por θ_{cam}. Assim, o vetor populacional das proporções nas respectivas categorias é representado por

$$\Theta = (\theta_{000}, \theta_{001}, \theta_{010}, \theta_{100}, \theta_{011}, \theta_{101}, \theta_{110}, \theta_{111}).$$

O vetor **Θ** é o alvo principal do estudo. É o parâmetro desconhecido e de interesse.

Com as restrições acima descritas, a função de verossimilhança pode ser escrita como:

$$L(\Theta \mid Y) = \Theta^{Y} = \prod_{c,a,m=0,1} \theta_{cam}^{y_{cam}}.$$

O produtório é sob o conjunto ***{(c,a,m); c,a,m = 0, 1}.***

A verossimilhança correspondente à amostra observada será então

$$L(\Theta \mid Y) = \left(\theta_{000}^{208} \times \theta_{001}^{79} \times \theta_{010}^{276} \times \theta_{100}^{91} \times \theta_{011}^{140} \times \theta_{101}^{15} \times \theta_{110}^{108} \times \theta_{111}^{36}\right)$$

O máximo dessa função – a estimativa de máxima verossimilhança – ocorre no ponto

$\hat{\theta}$ = (208, 79, 276, 91, 140, 15, 108, 36) $\dfrac{1}{953}$

= (0,2183; 0,083; 0,2896; 0,0955; 0,1469; 0,0157; 0,1133; 0,0378)

Tabela 1. Freqüências observadas

C	A	M	Θ	Y	Freqüência
0	0	0	θ_{000}	y_{000}	208
0	0	1	θ_{100}	y_{100}	79
0	1	0	θ_{010}	Y_{010}	276
1	0	0	θ_{001}	y_{001}	91
0	1	1	θ_{110}	y_{110}	140
1	0	1	θ_{011}	Y_{011}	15
1	1	0	θ_{101}	y_{101}	108
1	1	1	θ_{111}	y_{111}	36

6 – Estimativa de Θ e de parâmetros de interesse

Como o tamanho da amostra não pode ser decidido a priori, o espaço amostral – o espaço de todas as freqüências possíveis – não é bem determinado.

Isto dificulta o uso das técnicas usuais da estatística clássica ou frequentista. Consequentemente, a opção pelo uso das técnicas Bayesianas de análise é naturalmente adequada. De fato, com a especificação da verossimilhança seu estudo e análise são simplificados.

Utilizando a função de verossimilhança correspondente a amostra e a partir de uma densidade *a priori* Dirichlet de ordem 8 e com vetor de parâmetros formado por 0,25 nas oito posições, encontra-se a densidade *a posteriori* que, utilizando-se o símbolo de proporcionalidade, pode ser escrita como:

$$f(\Theta | Y) \propto \left(\theta_{000}^{207,25} \times \theta_{001}^{78,25} \times \theta_{010}^{275,25} \times \theta_{100}^{90,25} \times \theta_{011}^{139,25} \times \theta_{101}^{14,25} \times \theta_{110}^{107,25} \times \theta_{111}^{35,25}\right)$$

Com esta densidade, respectivamente, a média e a moda a posteriori de Θ são os seguintes vetores:

$$\tilde{\Theta} = \left(\frac{1}{955}\right) \times (208{,}25;\ 79{,}25;\ 276{,}25;\ 91{,}25;\ 140{,}25;\ 15{,}25;\ 108{,}25;\ 36{,}25) =$$
$$= (0{,}2181;\ 0{,}0830;\ 0{,}2893;\ 0{,}0955;\ 0{,}1469;\ 0{,}0160;\ 0{,}1134;\ 0{,}0380)$$

$$\hat{\Theta} = \left(\frac{1}{947}\right) \times (207{,}25;\ 78{,}25;\ 275{,}25;\ 90{,}25;\ 139{,}25;\ 14{,}25;\ 107{,}25;\ 35{,}25) =$$
$$= (0{,}2188;\ 0{,}0826;\ 0{,}2907;\ 0{,}0953;\ 0{,}1470;\ 0{,}0150;\ 0{,}1133;\ 0{,}0372)$$

Com a seguinte fórmula, calcula-se facilmente a matriz de covariância do vetor Θ:

$$Cov(\Theta | Y) = \left(\frac{1}{956}\right)\left(diag(\tilde{\Theta}) - \tilde{\Theta} \otimes \tilde{\Theta}^t\right)$$

Aqui, se A é um vetor de 8 componentes, a notação diag(A) representa a matriz quadrada de ordem 8, cujos elementos da diagonal são os componentes de A e os outros elementos da matriz são iguais a zero. O vetor coluna, formado pelos componentes de A, é representado por A^t. O símbolo $\otimes$ representa o produto matricial de dois vetores ou matrizes (o número de colunas do primeiro deve ser igual ao número de linhas do segundo).

Com o exposto acima, a conclusão é a de que podem ser calculadas todas as características desejáveis para descrever as informações sobre Θ que os dados oferecem. Conforme visto, a distribuição *a posteriori* de probabilidades de Θ é totalmente especificada. Para facilitar o entendimento das inferências, os cálculos foram realizados por etapas. Consideremos, inicialmente, a proporção populacional das aves sem algum dos 3 parasitas, θ_{000}. Esse parâmetro tem distribuição marginal Beta com parâmetros

a = 208,25 e b = 955 - 208,25 = 746,75.

A média e o desvio padrão serão, respectivamente,

$\boldsymbol{M_{000} = 0{,}2181}$ **e** $\boldsymbol{D_{000} = 0{,}0134}$**.**

Usando-se a distribuição beta com esses parâmetros ou a aproximação para uma normal com mesma média e mesmo desvio padrão, o intervalo com 95% de credibilidade para θ_{000} é (0,1919; 0,2442). Isto é, a proporção populacional de aves saudáveis, com alta probabilidade, é um número pertencente a este intervalo.

Considerando agora a população das aves com pelo menos um dos 3 tipos de parasitas. O parâmetro de interesse agora é um vetor de 7 componentes. Os componentes serão definidos como

$$\lambda_{cam} = \frac{\theta_{cam}}{1-\theta_{000}}$$

Dessa forma a soma dos componentes do novo vetor será igual à unidade. Foquemos agora a atenção na proporção das aves que possuem todos os 3 parasitas dentro da população das aves com pelo menos um parasita, isto é, o parâmetro $\lambda_{111} = \theta_{111}/(1-\theta_{000})$. Novamente, este parâmetro é distribuído segundo uma Beta, mas agora com parâmetros a = 36,25 e b = 710,5. A média e o Desvio padrão desse parâmetro são, respectivamente, os seguintes:

$$M_{111}^{*} = 0{,}04854 \quad \& \quad D_{111}^{*} = 0{,}00786$$

Com esta distribuição, obtemos o intervalo de 95% de credibilidade para λ_{111}: (0,03337; 0,06395).

Para o restante da população, a proporção de unidades com apenas um ou com dois tipos de parasitas é representada por $\delta = 1 - \theta_{000} - \theta_{111}$. As seguintes quantidades são agora o foco das inferências:

$$\Pi = \frac{1}{\delta}\left(\theta_{001}, \theta_{010}, \theta_{100}, \theta_{011}, \theta_{101}, \theta_{110}\right) \quad \& \quad \pi = \frac{\theta_{001} + \theta_{010} + \theta_{100}}{\delta}$$

O parâmetro Π é o vetor das proporções da população das aves com apenas um ou dois parasitas e π é a proporção total das aves com apenas um parasita nesta população. O vetor Π tem distribuição de Dirichlet com vetor de parâmetros igual à (79,25; 275, 25, 91,25; 140,25; 15,25; 108,25). Já o parâmetro π tem distribuição Beta com parâmetros a = 446,75 e b = 263,75. O intervalo com 95% de credibilidade para π é facilmente obtido: (0,5933; 0,6643).

Completando as estimações de interesse, o vetor Π é agora particionado no vetor de proporções de aves com apenas um parasita e no vetor de proporções de aves com exatamente dois tipos de parasitas. Denotemos esses dois vetores por:

$$\Sigma = \left(\sigma_{001}; \sigma_{010}; \sigma_{100}\right) = \left(\frac{1}{\theta_{001} + \theta_{010} + \theta_{100}}\right)\left(\theta_{001}; \theta_{010}; \theta_{100}\right) \quad \&$$

$$\Omega = \left(\omega_{011}; \omega_{101}; \omega_{110}\right) = \left(\frac{1}{\theta_{011} + \theta_{101} + \theta_{110}}\right)\left(\theta_{011}; \theta_{101}; \theta_{110}\right).$$

Ambos os vetores são distribuídos como Dirichlet, respectivamente, com vetores de parâmetros (79,25; 276,25; 91,25) e (140,25; 15,25; 108,25).

Lembremos que a soma dos elementos, tanto do vetor $\mathbf{\Sigma}$ quanto do vetor $\mathbf{\Omega}$, é igual à unidade. Assim, a região de credibilidade a ser construída é de dimensão 2 embora os vetores sejam de ordem 3. Por razões apenas estéticas foram escolhidos os pares $(\sigma_{010};\sigma_{100})$ e $(\omega_{011};\omega_{110})$ para a definição dos conjuntos de credibilidade das Figuras 1 e 2.

Figura 1. Região de credibilidade de 95% para $(\sigma_{010};\sigma_{100})$.

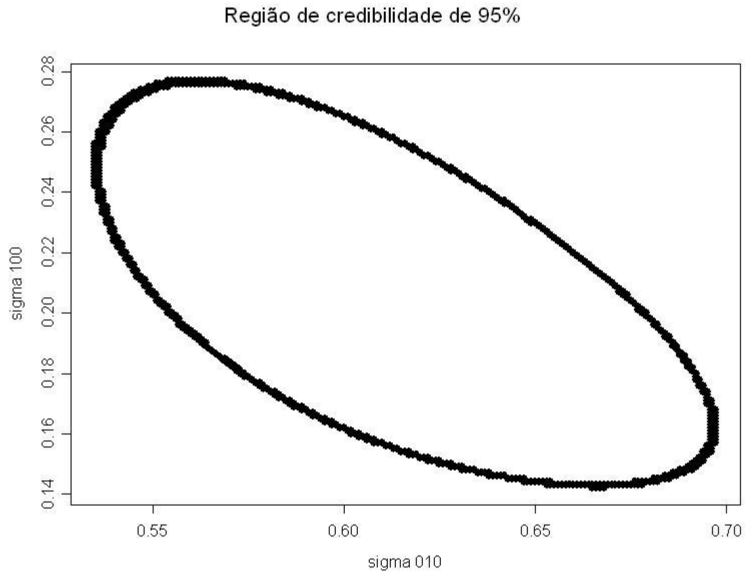
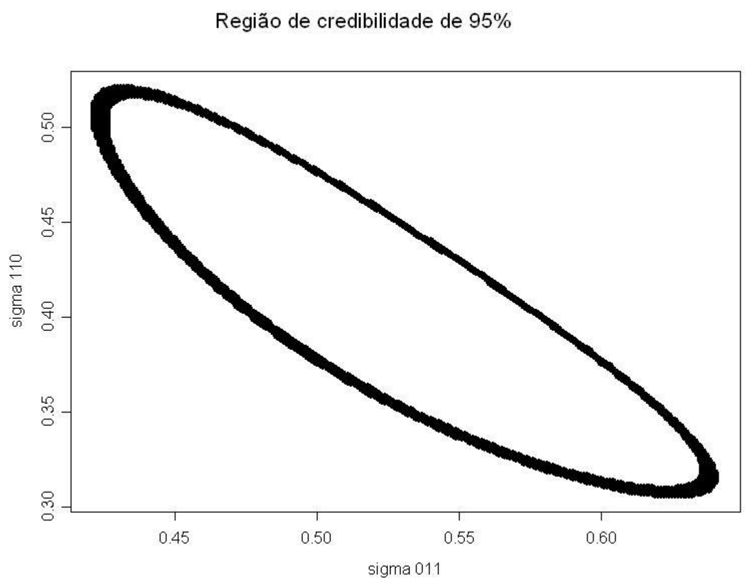


Figura 2. Região de credibilidade de 95% para $(\omega_{011};\omega_{110})$.



É natural que um leitor menos cuidadoso pergunte se não houve exagero teórico, ao considerar todo esse processo hierárquico para estimação dos parâmetros. Por que não considerar os intervalos de credibilidade para os parâmetros originais do vetor Θ? A razão para tal atitude é o fato de os elementos de Θ estarem associados negativamente. Isto limita a realização de um trabalho exclusivamente nas distribuições marginais de Θ. No entanto, como as regiões de credibilidade ilustram esse problema, foram apresentados na Tabela 2 os intervalos marginais para todos os componentes de Θ. Pode-se notar que a soma dos limites superiores dos intervalos irão produzir um número maior do que a unidade o que é um absurdo, pois a soma das proporções das diversas classificações somam 1.

Tabela 2: Intervalos de credibilidade, marginais, utilizando as marginais Betas com parâmetros α e β. LimInf e LimSup são os limites Inferior e superior dos intervalos.

Θ	α	β	Média	D. Padrão	LimInf	LimSup	Credibilidade
θ_{000}	208,25	746,75	0,218	0,013	0,192	0,244	95,02%
θ_{001}	79,25	875,75	0,083	0,009	0,065	0,100	95,06%
θ_{010}	276,25	678,75	0,289	0,015	0,261	0,318	95,01%
θ_{100}	91,25	863,75	0,096	0,010	0,077	0,114	95,05%
θ_{011}	140,25	814,75	0,147	0,011	0,124	0,169	95,03%
θ_{101}	15,25	939,75	0,016	0,004	0,008	0,024	95,35%
θ_{110}	108,25	846,75	0,113	0,010	0,093	0,133	95,04%
θ_{111}	36,25	918,75	0,038	0,006	0,026	0,050	95,15%

Esse trabalho termina com um teste de significância. O interesse é procurar saber qual dos parasitas apresenta a maior incidência (prevalência) na população de aves. Os parâmetros de interesse neste caso são as proporções de aves que apresentam cada uma das parasitas. Isto é, os parâmetros a serem comparados são os seguintes:

$$\theta_{1\bullet\bullet} = \theta_{111} + \theta_{110} + \theta_{101} + \theta_{100} \approx \quad Beta(251;704),$$
$$\theta_{\bullet 1\bullet} = \theta_{111} + \theta_{110} + \theta_{011} + \theta_{010} \approx \quad Beta(561;394)$$
$$\theta_{\bullet\bullet 1} = \theta_{111} + \theta_{011} + \theta_{101} + \theta_{001} \approx \quad Beta(271;684).$$

As hipóteses envolvendo estes parâmetros podem ser definidas como

$$
\begin{aligned}
H_{CA}: \theta_{1\bullet\bullet} < \theta_{\bullet 1 \bullet} \quad &\Leftrightarrow \quad H_{CA}: \theta_{101} + \theta_{100} < \theta_{011} + \theta_{010},\\
H_{AM}: \theta_{\bullet\bullet 1} < \theta_{\bullet 1 \bullet} \quad &\Leftrightarrow \quad H_{AM}: \theta_{101} + \theta_{001} < \theta_{110} + \theta_{010}\\
H_{CM}: \theta_{\bullet\bullet 1} < \theta_{\bullet 1 \bullet} \quad &\Leftrightarrow \quad H_{CM}: \theta_{011} + \theta_{001} < \theta_{110} + \theta_{100}.
\end{aligned}
$$

Para tornar simples o cálculo das probabilidades de cada uma destas hipóteses, os seguintes parâmetros foram definidos:

$$
\begin{aligned}
\Pi_{CA} &= \frac{\theta_{101} + \theta_{100}}{\theta_{101} + \theta_{100} + \theta_{011} + \theta_{010}} \quad \approx \quad Beta(106{,}5;416{,}5);\\
\Pi_{AM} &= \frac{\theta_{101} + \theta_{001}}{\theta_{101} + \theta_{001} + \theta_{110} + \theta_{010}} \quad \approx \quad Beta(94{,}5;384{,}5) \quad \&\\
\Pi_{CM} &= \frac{\theta_{011} + \theta_{001}}{\theta_{011} + \theta_{001} + \theta_{110} + \theta_{100}} \quad \approx \quad Beta(219{,}5;199{,}5).
\end{aligned}
$$

O leitor deve entender que o cálculo das probabilidades de cada uma das hipóteses é equivalente ao cálculo das probabilidades destes últimos parâmetros serem menores do que 0,5. Note que para estes cálculos foram considerados vetores divididos pela soma de seus elementos. Como antes, foram produzidas outras distribuições de Dirichlet ou distribuições Betas. Os cálculos necessários para tomadas de decisão sobre as hipóteses estão apresentados a seguir.

$$
\begin{aligned}
\Pr\{H_{CA}\} &= \Pr\{\Pi_{CA} < \tfrac{1}{2}\} = 1,\\
\Pr\{H_{AM}\} &= \Pr\{\Pi_{AM} < \tfrac{1}{2}\} = 1\\
\Pr\{H_{CM}\} &= \Pr\{\Pi_{CM} < \tfrac{1}{2}\} = 0{,}164,
\end{aligned}
$$

A conclusão é a de que os Ácaros são as parasitas com maior prevalência entre os estudados. Os carrapatos são aqueles com menor prevalência. Para alguém acostumado a usar testes de significância, pode-se utilizar a alternativa Bayesiana para o p-valor, o e-valor, que é um cálculo de significância no espaço paramétrico. Sua descrição pode ser encontrada em Pereira & Stern (1999) e corresponde ao cálculo da probabilidade, na distribuição a posteriori, do conjunto de pontos que possuem densidade inferior ao ponto que define a hipótese. Testando se as prevalências são iguais, foram encontrados os seguintes e-valores.

1 – Hipótese 1: Carrapatos e Ácaros possuem a mesma prevalência:

Evidência (C=A) = 0,004%.

2 – Hipótese 2: Ácaros e Malófagos possuem a mesma prevalência:

Evidência (A=M) = 0,004%

3 – Hipótese 3: Carrapatos e Malófagos possuem a mesma prevalência:

Evidência (C=M) = 0,005%

Com estes resultados pode-se afirmar que as prevalências são diferentes. A ordem da maior para a menor é a seguinte: Ácaros, Malófagos e Carrapatos. Completando o conjunto de técnicas estatísticas Bayesianas utilizadas, a Tabela 3 apresenta os intervalos de credibilidade para as proporções de incidência de cada um dos tipos parasitas.

Tabela 3: Intervalos de credibilidade para a incidência dos 3 tipos de parasitas, utilizando novamente Betas com parâmetros α e β. LimInf e LimSup são os limites Inferior e superior dos intervalos.

Parâmetro	α	β	Média	D. Padrão	LimInf	LimSup	Credibilidade
$\theta_{1\bullet\bullet}$	251	704	0,26	0,014	0,235	0,291	95%
$\theta_{\bullet 1\bullet}$	561	394	0,59	0,016	0,556	0,619	95%
$\theta_{\bullet\bullet 1}$	271	684	0,28	0,015	0,255	0,312	95%

7 – Referências Bibliográficas

CUNNINGHAM, A.(1995). **Disease risks of wildlife translocations**. Conservation Biology 10, pág.349-353.

FLECHTMANN, C. H. W.(1990). **Ácaros de importância médico veterinária.** Nobel, São Paulo.

PEREIRA, C. A, de B. E STERN, J. M. (2005). **Inferência indutiva com dados discretos: uma visão genuinamente Bayesiana.** Antofagasta : Universidad de Antofagasta (XV Congreso de Matemática Capricórnio). 104p.

PEREIRA C. A. de B., STERN, J. M. (1999). **Evidence and Credibility: Full Bayesian Significance Test for Precise Hypotheses.** *Entropy Journal***,** I:104-115. (http://www.mdpi.org/entropy/).