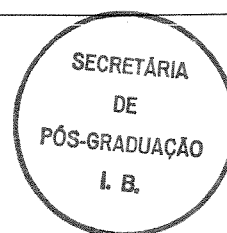


Universidade Estadual de Campinas
Instituto de Biologia



UM MODELO PARA A ESTIMAÇÃO
DO NÚMERO DE TARTARUGAS MARINHAS
DESOVANDO EM UMA PRAIA EM UMA TEMPORADA

Paulo Cesar Rosito Barata

Orientador: Dr. Ulisses Caramaschi

Tese apresentada ao Instituto de Biologia
da Universidade Estadual de Campinas
como parte dos requisitos para a obtenção do título de
Doutor em Ciências (Ecologia)

Este exemplar corresponde à redação final
da tese do(a) candidato(a)
Paulo Cesar Rosito Barata
e aprovada pela Comissão Julgadora.
22/12/96 *U. Caramaschi*

B231m

29466/BC

Campinas

1996

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA CENTRAL DA UNICAMP

B231m

Barata, Paulo Cesar Rosito

Um modelo para a estimação do número de tartarugas
marinhas desovando em uma praia em uma temporada /
Paulo Cesar Rosito Barata . - - Campinas, SP : [s.n.], 1996.

Orientador : Ulisses Caramaschi..

Tese (doutorado) - Universidade Estadual de
Campinas, Instituto de Biologia.

1. Ecologia. 2. Demografia. 3. Animais - População.
4. Tartaruga marinha. 5. Modelos matemáticos. 6. Simu-
lação. 7. Estimativas. I. Caramaschi, Ulisses.
II. Universidade Estadual de Campinas. Instituto de
Biologia. III. Título.

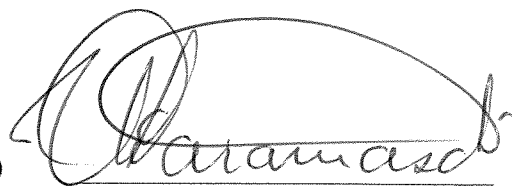
UNIDADE	BC
N.º CHAMADA:	T/Unicamp
	B231m
V.	Ex.
TOMBO BC/	29466
PROC.	281/97
C	☐
D	☒
PREÇO	R\$ 11,00
DATA	17/01/97
N.º CPD	

Campinas, 02 de dezembro de 1996

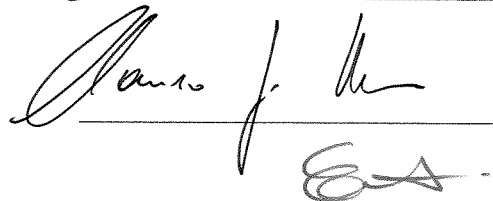
BANCA EXAMINADORA:

TITULARES:

Prof. Dr. Ulisses Caramaschi (orientador)



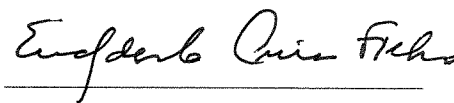
Prof. Dr. Cláudio José Struchiner



Prof. Dr. Eduardo Massad



Prof. Dr. Euclydes Custódio de Lima Filho



Prof. Dr. Woodruff Whitman Benson



SUPLENTE:

Prof. Dr. Flávio Antônio Mães dos Santos



Prof. Dr. Rodney Carlos Bassanezi



Dedico este trabalho

ao Professor George Svetlichny,
um amigo de tantos anos, meu mestre,

e à Aninha,
que teve o maior cuidado do mundo com esta minha tese.

Índice

Agradecimentos vi

Resumo x

Summary xii

1 Introdução 1

2 Sumário da biologia reprodutiva das tartarugas marinhas *Caretta caretta* 9

2.1. Biologia reprodutiva 10

3 Métodos computacionais e dados utilizados nas simulações 15

3.1. Computadores e linguagens de programação 15

3.2. Rotinas-chave nas simulações 16

3.3. Distribuições-modelo do número de ninhos por tartaruga 16

3.4. Construção dos arquivos de dados usados nas simulações 19

4 Uma análise do problema 23

4.1. O processo de observação na praia 24

4.2. Metodologia de simulação 25

4.3. O processo de observação na praia: resultados 27

4.4. O processo de observação na praia: discussão 27

4.5. A transformação da distribuição de probabilidade 29

5 O modelo 32

5.1. Definição do estimador 32

5.2. Existência e unicidade da solução 37

5.3. Discussão 38

Apêndice 5.A. Dedução da equação (5.7) 39

6	Propriedades estatísticas do modelo	42
6.1.	Metodologia de simulação	42
6.2.	Metodologia de análise dos resultados das simulações - a região útil do estimador	47
6.3.	Resultados das simulações: propriedades do estimador	50
6.4.	Cálculo do esforço de amostragem	55
6.5.	Discussão	58
7	Análise de sensibilidade	59
7.1.	Sensibilidade em relação a ω_e : metodologia de simulação	59
7.2.	Sensibilidade em relação a ω_e : resultados	65
7.3.	O problema da perda de marcas	65
8	Construção de um intervalo de confiança	69
8.1.	O intervalo de confiança	69
8.2.	Discussão sobre o algoritmo	73
8.3.	Avaliação do <i>bootstrap</i> : metodologia de simulação	74
8.4.	Calibração do intervalo de confiança	76
8.5.	Avaliação do <i>bootstrap</i> : resultados	80
8.6.	Discussão	80
9	Teste do modelo	84
9.1.	O teste	84
9.2.	Avaliação do teste: metodologia de simulação	87
9.3.	Calibração do teste	88
9.4.	Avaliação do teste: resultados	89
9.5.	Discussão	91
10	Exemplo: aplicação do modelo a dados das tartarugas marinhas <i>Caretta caretta</i> da Praia do Forte, Bahia	92
10.1.	Metodologia de campo do Projeto TAMAR	92
10.2.	Metodologia aplicada aos dados da Praia do Forte	96
10.3.	Os dados da Praia do Forte: resultados e discussão	99

10.4. Aplicação do modelo aos dados da Praia do Forte: metodologia	102
10.5. Aplicação do modelo aos dados da Praia do Forte: resultados e discussão	111
11 Conclusões	114
Apêndice A	116
Apêndice B	120
Apêndice C	124
Referências bibliográficas	128

Agradecimentos

Amigos que há muito tempo eu não encontrava, outros com quem convivo com frequência. Dúvidas respondidas de longe, via Internet, ou esclarecidas aqui mesmo, no Rio de Janeiro. Instituições que me abriram suas portas para possibilitar a realização desta tese. Por todos fui sempre recebido da melhor forma possível. A todos, minha gratidão pela imensa ajuda que tive ao longo do meu trabalho.

Quero agradecer, inicialmente, a três pessoas que foram centrais na realização desta pesquisa:

Ao Professor Ulisses Caramaschi, do Museu Nacional, Universidade Federal do Rio de Janeiro (UFRJ), pela extrema atenção que dele recebi ao longo do meu doutorado, pelas muitas e valiosas discussões sobre herpetologia e por toda a orientação que tive na realização desta tese.

Ao Professor George Svetlichny, do Departamento de Matemática da Pontifícia Universidade Católica do Rio de Janeiro (PUC-RJ), com quem tive inúmeras discussões que me ajudaram a formular o modelo do Capítulo 5, que é o fundamento desta tese, e compreender suas propriedades. Agradeço especialmente ao Professor George por ter aceito trazer tantas horas de seu tempo ao estudo de um tema (tartarugas marinhas!) tão distante de seus assuntos de trabalho.

À oceanóloga Neca Marcovaldi, Presidente da Fundação Pró-TAMAR (Projeto TAMAR / IBAMA), Praia do Forte, Bahia, que, além de colocar à minha disposição o banco de dados do Projeto TAMAR e material bibliográfico e fazer uma revisão de trechos do Capítulo 10, me transmitiu muito de sua experiência em relação às tartarugas marinhas e me deu sempre o máximo incentivo nesta pesquisa.

O banco de dados do Projeto TAMAR, que utilizei não apenas nos cálculos do Capítulo 10 mas como fonte de idéias ao longo de toda esta tese, é resultado de muitos anos de trabalho tartarugueiro, realizado por um grande número de pessoas em muitas praias de diversos estados do Brasil, não somente na Praia do Forte. Assim, além de Neca, quero agradecer também a Guy Marcovaldi (Coordenador Nacional do Projeto TAMAR / IBAMA), Cláudio, Taisi, Luciana, Cláudia, Cecília, Joca, Alex Silveira, Juarez, Toninho, Jaqueline, Cesar, Júlio, Eduardo, Adriana, Mônica, Nêgo, Vitor, Eron, Seu Geraldo, Cheiro, Miúdo, Gil, George, Roberto, Paulo Lara, Badanha, Berenice, Hugo, Alex Santos, Deborah, Márcia, Denise e a tantos outros cujo esforço e dedicação na conservação das tartarugas marinhas estão contidos nestes dados.

Agradeço também a todas as demais pessoas e instituições que tornaram possível este trabalho:

Ao Professor Cláudio José Struchiner, da Fundação Oswaldo Cruz (Fiocruz), Rio de Janeiro, por todas as discussões que tivemos sobre diversos tópicos de estatística relacionados a esta tese, pela leitura de trechos deste trabalho em sua fase inicial e pela atenção com que sempre me atendeu.

Ao Dr. Sérgio Granville, do Centro de Pesquisas de Energia Elétrica (CEPEL / Eletrobrás), Rio de Janeiro, pela discussão inicial sobre os meandros dos métodos numéricos para a solução de problemas de otimização.

Ao Professor Geovan Tavares, do Departamento de Matemática da PUC-RJ, pela atenção em relação às minhas questões sobre geometria.

Ao Professor Bradley Efron, da Universidade de Stanford, USA, que, através da Internet, me esclareceu um ponto essencial sobre a utilização do método do *bootstrap*.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa de estudos concedida para a realização do meu doutorado, processo nº. 142204 / 92-7.

Ao Laboratório Nacional de Computação Científica (LNCC / CNPq), Rio de Janeiro, que me colocou à disposição o seu sistema de computação, no qual realizei as simulações principais desta tese. Agradeço especialmente a Flávio Barbosa Toledo e Sérgio Ricardo Alves de Souza, da Divisão de Suporte, e a Lino Oliveira Sobral e Alfredo Veiga de Carvalho, do Serviço de Apoio ao Usuário, pela forma amigável como fui recebido e por toda a assistência que tive no meu trabalho. Agradeço também à Biblioteca do LNCC pela utilização do seu acervo nas áreas de computação e matemática aplicada.

À Biblioteca Central da Fundação Oswaldo Cruz, Rio de Janeiro, pela utilização de seu acervo na área de ciências biológicas e por contatos com outras bibliotecas do Brasil e do exterior. Agradeço aos bibliotecários Marilene Fragas Costa, Glécio Batista da Silva e Mauro Crispim da Silva pela ajuda na obtenção de material bibliográfico e reprodução de trabalhos.

À Biblioteca do Instituto de Matemática Pura e Aplicada (IMPA / CNPq), Rio de Janeiro, pela utilização de seu acervo nas áreas de matemática, estatística e computação. Agradeço aos bibliotecários Maria da Graça Araújo Penna Duarte e Sérgio de Souza Telles e à equipe da biblioteca pela atenção que sempre me deram.

À Biblioteca do Museu Nacional (UFRJ), Rio de Janeiro, pela utilização de seu acervo na área de ciências biológicas. Agradeço aos bibliotecários José Arimatéia de Aguiar e Antônio Carlos Gomes Lima pela ajuda na obtenção de material bibliográfico.

Ao Archie Carr Center for Sea Turtle Research, Universidade da Flórida, USA, pela utilização da "Sea Turtle Bibliographic Database", que este centro de pesquisas torna disponível para consultas através da Internet.

Aos vários pesquisadores - peço desculpas por não citar nomes - que me enviaram, do Brasil e do exterior, cópias de trabalhos necessários à minha pesquisa.

À Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz, Rio de Janeiro, pelo tempo precioso que me concedeu para a realização do meu doutorado.

À Coordenação do Programa de Pós-Graduação em Ecologia, Instituto de Biologia, Universidade Estadual de Campinas (Unicamp), por todo o apoio durante o meu doutorado.

A Carlos Osanai, Miriam Struchiner, Cláudia Travassos, Célia Szwarcwald, Martha Argel e Deborah Faria, meus colegas da Fiocruz, UFRJ e Unicamp, pela grande ajuda que me deram no uso de editores de texto, confecção de gráficos, correspondências e muitas outras coisas necessárias a uma tese.

A Mariana Rocha e Adriana Carvalho, que desenharam no computador as figuras do Capítulo 5.

A Marcia Chame, minha colega da Fiocruz e de caminho no doutorado, pelas mil e uma conversas sobre mil e uma questões relativas a esta tese e à ecologia.

À minha irmã Cristina, que atendeu sempre com muito boa vontade a todos os meus 9237 telefonemas com dúvidas sobre questões de linguagem.

Por fim, quero agradecer a todos os meus amigos - foram muitos - pelo carinho e pela presença no decorrer deste trabalho.

Rio de Janeiro, outubro de 1996

Resumo

Um modelo para a estimação do número de tartarugas marinhas desovando em uma praia em uma temporada

Nesta tese é apresentado um modelo para a estimação do número de tartarugas marinhas desovando em uma praia em uma temporada de desova, tendo por base dados obtidos por marcação das tartarugas na praia de desova e a contagem do número total de ninhos depositados na praia na temporada.

O estimador classicamente utilizado neste problema calcula o número de fêmeas desovando na temporada como sendo a razão entre o número total de ninhos depositados na praia na temporada e o número médio de ninhos depositados por tartaruga na temporada. Caso a cobertura do trabalho de marcação na praia não seja completa, o número médio de ninhos depositados por tartaruga na temporada, conforme calculado diretamente dos dados de campo, será em geral menor do que o número médio real. Este fato, que já foi reconhecido por diversos pesquisadores, faz com que o estimador clássico, caso o número total de ninhos depositados na praia na temporada seja conhecido, produza uma estimativa para o número de tartarugas desovando na temporada em geral maior que o número real.

O procedimento aqui proposto aplica uma correção ao estimador clássico mencionado acima, utilizando, no lugar do número médio de ninhos por tartaruga

calculado diretamente dos dados obtidos por marcação, um número médio de ninhos estimado através de uma modelagem matemática do processo de observação das tartarugas na praia. A essência do modelo é a análise, feita com o emprego de técnicas de probabilidade, da transformação existente entre a distribuição de probabilidade real do número de ninhos depositados por tartaruga na temporada e a distribuição de probabilidade do número de ninhos depositados por tartaruga na temporada calculada apenas para as tartarugas observadas pela equipe de campo na temporada, na situação em que nem todas as tartarugas tenham sido encontradas por ocasião da desova.

O modelo tem suas propriedades estatísticas investigadas por meio de simulações em computador. Os resultados destas simulações mostram que o estimador proposto é essencialmente não-viciado (tem vício relativamente pequeno) e consistente. As simulações permitem também uma estimativa do esforço de trabalho de campo necessário para a obtenção de uma precisão especificada na estimação do número de tartarugas desovando na temporada.

São propostos também (1) um procedimento para a construção de um intervalo de confiança para a população estimada, onde é utilizado o método do *bootstrap*, um dos chamados métodos de reamostragem, e (2) um teste para o modelo; ambos têm sua validade investigada também por simulações em computador. Como exemplo, o modelo é aplicado às tartarugas marinhas *Caretta caretta* desovando na Praia do Forte, estado da Bahia, utilizando-se dados colhidos pelo Projeto Tartaruga Marinha (Projeto TAMAR), do Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA). Por fim, algumas possíveis extensões do modelo são sugeridas.

Summary

A model for the estimation of the number of sea turtles nesting on a beach in a season

In this thesis, a model for the estimation of the number of sea turtles nesting on a beach in a season is presented, based on data gathered by tagging the turtles on the nesting beach and counting the total number of nests laid on the beach in the season.

The estimator classically used in this kind of problem calculates the number of females nesting in the season as the ratio between the total number of nests laid on the beach in the season and the mean number of nests laid per turtle in the season. If the beach tagging coverage is not complete, the mean number of nests laid per turtle in the season directly calculated from the beach data will be generally smaller than the actual mean number. As a result of this fact, already recognized by several researchers, the classical estimator, when the total number of nests laid on the beach in the season is known, produces an estimate for the number of turtles nesting in the season generally greater than the actual number of turtles.

The procedure proposed here applies a correction to the classical estimator mentioned above, using, instead of the mean number of nests per turtle directly calculated from the data gathered through tagging, a mean number of nests

estimated through a mathematical model of the sampling of the turtles on the beach. The essence of the model is the analysis, using the probability theory, of the transformation occurring between the actual probability distribution of the number of nests laid per turtle in the season and the probability distribution of the number of nests laid per turtle in the season calculated only for the turtles found by the tagging team in the season, whenever the beach tagging coverage is not complete.

The model has its statistical properties investigated by means of computer simulations. The results of these simulations show that the proposed estimator is essentially unbiased (it has a relatively small bias) and consistent. The simulations also allow an estimate of the beach tagging coverage necessary to obtain a specified precision in the estimation of the number of turtles nesting in the season.

A procedure for the construction of a confidence interval for the estimated population, using the bootstrap method, and a test of the model are also proposed. Both have their validity checked by means of computer simulations. As an example, the model is applied to the *Caretta caretta* sea turtles nesting on Praia do Forte, Bahia, Brazil, using data gathered by Projeto Tartaruga Marinha (TAMAR Project / IBAMA), the Brazilian sea turtle conservation program. Finally, some possible extensions of the model are proposed.

CAPÍTULO 1

Introdução

O tamanho de uma população é freqüentemente um parâmetro de importância central no estudo da ecologia de populações animais (Krebs, 1989; Begon *et al.*, 1990). Na teoria da ecologia, uma grande classe de modelos para a explicação de fenômenos fundamentais, particularmente no estudo de interações entre espécies, tem entre seus parâmetros o tamanho das populações estudadas (May, 1981; Emlen, 1984). Na ecologia aplicada, a estimação¹ de tamanho de população é um problema que se apresenta em diversas áreas, como por exemplo na modelagem da dinâmica de populações (Johnson, 1994), monitoramento de populações em risco de extinção (Dennis *et al.*, 1991) e determinação de estoques para fins de possível exploração comercial (Raftery e Schweder, 1993). Na biologia de conservação, a avaliação do tamanho de populações é reconhecida como uma das informações básicas no acompanhamento do *status* de uma espécie (McNeely *et al.*, 1990).

Em relação às tartarugas marinhas, a estimação do tamanho de suas populações é um tópico que tem atraído a atenção de diversos pesquisadores

¹ Neste trabalho, conforme convenção usual em estatística, o ato, o procedimento de estimar um parâmetro de uma população é denominado de *estimação*. Um *estimador* é uma função que, a partir de dados observados, produz uma aproximação para o parâmetro sendo estimado, aproximação esta que é denominada de *estimativa*.

(Hughes, 1974a,b; Schulz, 1975; Carr *et al.*, 1978; Meylan, 1982; Mrosovsky, 1983; Limpus, 1985, 1992a; Le Gall *et al.*, 1986; Mortimer e Carr, 1987; Mortimer, 1988; Seabrook, 1991), sendo uma questão de reconhecida relevância no planejamento e avaliação de atividades de conservação (Carr, 1980; Meylan, 1982). Uma motivação importante para o desenvolvimento de métodos de estimação de população é a construção, validação e estimação de parâmetros de modelos populacionais para as tartarugas marinhas (Richardson, 1982; Frazer, 1986; Crouse *et al.*, 1987). O plano global para a conservação das tartarugas marinhas elaborado pelo grupo de especialistas em tartarugas marinhas da IUCN (The World Conservation Union) coloca a estimação do tamanho de populações e a modelagem populacional em geral como áreas prioritárias de pesquisa no tocante à conservação destes animais (IUCN/SSC, 1995).

No caso das tartarugas marinhas, diversos problemas, alguns dos quais são listados abaixo, se compõem para tornar a tarefa de estimar o tamanho de uma população, e de forma geral o estudo da dinâmica populacional, especialmente difíceis (Carr *et al.*, 1978; Carr, 1980; Pritchard, 1980; Meylan, 1982; Bowen *et al.*, 1992; Limpus, 1994). Os pontos descritos a seguir, em essência já apresentados por Meylan (1982), servem apenas para exemplificar as dificuldades encontradas no estudo de populações destas tartarugas, sem haver aqui a intenção de se elaborar um quadro completo sobre sua ecologia de populações. Ademais, estes pontos não se aplicam necessariamente a todas as sete espécies de tartarugas marinhas reconhecidas atualmente (Pritchard, 1979), sendo no entanto em geral válidos para as tartarugas *Caretta caretta* e *Chelonia mydas*, que são as duas espécies das quais melhor se conhece atualmente a dinâmica populacional.

1. Uma população de tartarugas marinhas em geral distribui-se por áreas muito extensas do mar, o que impede a realização de censos e dificulta sua amostragem (Carr, 1980; Meylan, 1982). Tartarugas marcadas em praias de desova podem ser encontradas, geralmente em áreas de alimentação, a mais de 2000 km de distância do local de marcação (Mortimer e Carr, 1987; Limpus *et al.*, 1992; Green, 1994).

2. As tartarugas marinhas tendem a desovar colonialmente, com uma mesma praia ou um conjunto de praias geograficamente delimitado congregando tartarugas provenientes de diversas áreas de alimentação, freqüentemente muito distantes entre si (Carr e Hirth, 1962; Carr *et al.*, 1978; Limpus *et al.*, 1992).

3. Pode existir uma superposição geográfica entre áreas de alimentação de tartarugas pertencentes a colônias distintas, desovando em praias bastante distantes entre si (Carr, 1975; Limpus *et al.*, 1992).

4. Indivíduos em diferentes estágios de crescimento ocupam diferentes habitats no mar, desde o habitat aparentemente pelágico dos filhotes recém-nascidos e tartarugas ainda relativamente pequenas (Witham, 1980; Carr, 1987; Bolten *et al.*, 1994) a habitats em geral costeiros das tartarugas juvenis e adultas (Carr, 1987; Limpus *et al.*, 1994).

5. As tartarugas marinhas na natureza têm um período de geração extremamente elevado, havendo para as *Caretta caretta* estimativas de idade de maturação reprodutiva da ordem de 35 anos ou mais (Foster, 1994; Frazer *et al.*, 1994) e para as *Chelonia mydas* estimativas de mais de 50 anos (Frazer e Ladner, 1986).

6. Pode existir grande variação no número de fêmeas desovando em uma mesma praia em anos consecutivos (Carr *et al.*, 1978; Richardson, 1982).

O item 1 acima indica a escala geográfica das pesquisas sobre a ecologia das tartarugas marinhas. Os itens 2 a 4 mostram a dificuldade de se relacionar as populações das áreas de alimentação com as populações desovando em praias determinadas, bem como a dificuldade de se relacionar populações de tartarugas em diferentes estágios de crescimento. Os itens 5 e 6 mostram que pesquisas em dinâmica populacional, no caso das tartarugas marinhas, requerem períodos de tempo consideráveis para a formação de uma base de dados sólida.

Uma característica das populações de tartarugas marinhas é a existência de colônias, populações com alto grau de isolamento genético umas em relação às outras. Para efeitos de estimação de população, da mesma forma que para efeitos de atividades de conservação, a colônia é em princípio uma unidade natural na delimitação de uma população (Carr, 1975; Bowen *et al.*, 1992; Bowen *et al.*, 1994). Contudo, obter uma estimativa para a população total de uma colônia, incluindo indivíduos em todos os estágios de crescimento, implicaria, além da dificuldade de definir exatamente cada colônia (Gyuris e Limpus, 1988), na necessidade de pesquisas em diversas áreas de alimentação geograficamente separadas, com todas as dificuldades operacionais trazidas pela amostragem de populações no mar (Bustard, 1979; Meylan, 1982). Além disso, nem todas as áreas de alimentação são necessariamente conhecidas (Meylan, 1982), sem contar que muito pouco se conhece sobre o habitat dos filhotes e tartarugas ainda relativamente pequenas (Witham, 1980; Carr, 1987; Bolten *et al.*, 1994).

Por estes motivos, embora atualmente já sejam encontrados exemplos de trabalhos estimando a população de tartarugas marinhas em áreas de alimentação, utilizando modelos de marcação e recaptura (Limpus, 1985, 1992a), a estimação do número de fêmeas reprodutoras a partir dos dados colhidos nas praias de desova é a escolha usual e na maioria das vezes a única viável no tocante à estimação de população destas tartarugas (Carr *et al.*, 1978; Meylan, 1982). Na realidade, a estimação do número de fêmeas reprodutoras referentes a uma determinada praia, ou a um conjunto de praias, é um índice do tamanho da população total (Meylan, 1982). Este índice poderia ser representado, de forma alternativa, pelo número de rastros ou de ninhos deixados na praia (Carr, 1980; Meylan, 1982), o que é feito por diversos programas de monitoramento de populações de tartarugas marinhas (Shoop *et al.*, 1985; Owen *et al.*, 1994; Schroeder, 1994). Entretanto, há situações em que o número de rastros ou de ninhos precisam ser convertidos para o número de tartarugas correspondentes, como por exemplo na elaboração de tabelas de vida (Frazer, 1986), ajuste e validação de modelos para a dinâmica populacional (Richardson, 1982) e avaliação de dados referentes à captura acidental de tartarugas marinhas (Hillestad *et al.*, 1982).

A estimação do número de fêmeas reprodutoras a partir de dados obtidos nas praias de desova pode ser decomposta em dois diferentes problemas, conforme já exposto por Meylan (1982):

1. Estimação do número de fêmeas desovando em uma temporada de desova.

2. Estimação do número total de fêmeas reprodutoras existentes na população, a partir do número de fêmeas desovando em cada temporada. Este problema existe devido ao fato de que nem todas as fêmeas reprodutoras da população desovam em uma mesma temporada.

A estimação do número de fêmeas que desovam em uma temporada, com base em dados obtidos nas praias de desova por meio da marcação das tartarugas, já foi realizada por vários pesquisadores (Hughes, 1974a; Schulz, 1975; Carr *et al.*, 1978; Marquez *et al.*, 1982; Le Gall *et al.*, 1986; Mortimer e Carr, 1987; Mortimer, 1988). No entanto, não existe na literatura um tratamento estatístico adequado deste problema. Por tratamento estatístico adequado entende-se:

(1) a definição de um bom estimador, uma função que produza, a partir dos dados disponíveis, uma estimativa pontual da população aceitável do ponto de vista da estatística e (2) uma medida da variabilidade amostral da estimativa, o que pode ser traduzido em um intervalo de confiança para a população sendo estimada.

No único trabalho publicado que procura aplicar um procedimento estatístico à estimação da população de fêmeas desovando em uma temporada, Le Gall *et al.* (1986) tentaram aplicar o modelo de captura e recaptura de Jolly-Seber para populações abertas (Pollock *et al.*, 1990) a dados obtidos por meio de marcação para tartarugas *Chelonia mydas* desovando em praias do Oceano Índico. No entanto, a aplicação tentada por estes pesquisadores, que procuram estimar o número de fêmeas desovando a cada dia em uma praia, claramente viola as premissas do modelo de captura e recaptura de Jolly-Seber para populações abertas, o que é, de forma incompleta, reconhecido por estes próprios autores.

Embora diferentes métodos possam ser encontrados na literatura para a estimação do número de fêmeas desovando em uma praia em uma temporada (Hughes, 1974a; Richardson *et al.*, 1978; Marquez *et al.*, 1982; Limpus, 1985), o método classicamente empregado neste problema de estimação calcula o número de tartarugas como sendo a razão entre o número total de ninhos depositados na praia na temporada e o número médio de ninhos depositados por tartaruga na temporada (Schulz, 1975; Le Gall *et al.*, 1986; Mortimer e Carr, 1987; Mortimer, 1988). Este método é de aplicação intuitiva, uma vez reconhecido que uma mesma tartaruga pode desovar mais de uma vez na temporada e se disponha de uma contagem, ou estimativa, do número total de ninhos. No entanto, como será discutido no Capítulo 4, o método, caso a cobertura do trabalho de marcação na praia não seja completa, produz uma estimativa para o número de tartarugas em geral maior que o número real, já que a cobertura incompleta do trabalho de marcação produz um número médio de ninhos por tartaruga em geral menor que o número médio real. Esta diferença entre a média observada do número de ninhos por tartaruga e a média real, caso nem todas as tartarugas sejam encontradas na praia no momento da desova, já foi reconhecida por diversos pesquisadores (Pritchard e Trebbau, 1984; Frazer e Richardson, 1985; Mortimer e Carr, 1987), mas não houve até o presente momento a proposta de uma solução para este problema.

Esta tese tem como objetivo apresentar um modelo para a estimação do número de fêmeas desovando em uma praia em uma temporada, com base em dados obtidos por marcação das tartarugas e pela observação do número total de desovas. Na realidade, em essência é aplicada uma correção ao estimador

clássico mencionado acima, utilizando-se, no lugar do número médio de ninhos por tartaruga calculado diretamente dos dados obtidos por marcação, um número médio de ninhos estimado através de uma modelagem matemática do processo de observação das tartarugas na praia. O modelo aqui proposto tem suas propriedades estatísticas investigadas por meio de simulações em computador. São propostos também um procedimento para a construção de um intervalo de confiança para a população estimada e um teste para o modelo, que têm sua validade investigada também por simulações em computador. Como exemplo, o modelo é aplicado às tartarugas marinhas *Caretta caretta* desovando na Praia do Forte, estado da Bahia, utilizando-se dados colhidos pelo Projeto Tartaruga Marinha (Projeto TAMAR), do Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA). Por fim, algumas possíveis extensões do modelo são sugeridas.

CAPÍTULO 2

Sumário da biologia reprodutiva das tartarugas marinhas *Caretta caretta*

Neste capítulo será apresentado um resumo da biologia reprodutiva das tartarugas marinhas *Caretta caretta* fêmeas, nos aspectos que sejam necessários para a compreensão das questões discutidas no presente trabalho. Para uma revisão geral sobre a biologia das tartarugas *Caretta caretta*, consulte Dodd (1988). Hirth (1980) e Ehrhart (1982) apresentam revisões sobre a reprodução das tartarugas marinhas, com dados para as diferentes espécies, incluindo as *Caretta caretta*. As teses de Richardson (1982) e Limpus (1985), além de dados sobre a biologia das tartarugas *Caretta caretta*, contêm discussões valiosas sobre o significado de diversos parâmetros relativos à sua reprodução e possíveis problemas na coleta de dados referentes a estes parâmetros.

Deve ser observado que a história natural das sete espécies de tartarugas marinhas, nos aspectos tratados neste resumo, é fundamentalmente similar (Carr *et al.*, 1978; Buskirk e Crowder, 1994): em condições naturais, todas as espécies desovam em geral mais de uma vez em uma temporada, produzem um número de ovos relativamente grande (entre 50 e 200 ovos em geral) a cada desova, têm um comportamento em relação à desova semelhante, usualmente (com exceção do

gênero *Lepidochelys*) não desovam em temporadas sucessivas e geralmente migram entre áreas de alimentação e de reprodução (Hirth, 1980; Witzell, 1983; Pritchard e Trebbau, 1984; Dodd, 1988; Limpus *et al.*, 1993; Reichart, 1993; Marquez, 1994). Pritchard e Trebbau (1984) apresentam uma revisão sobre a biologia das cinco espécies de tartarugas marinhas encontradas na América do Sul: *Caretta caretta*, *Chelonia mydas*, *Eretmochelys imbricata*, *Lepidochelys olivacea* e *Dermochelys coriacea*.

Na discussão que se segue, todos os parâmetros (ex. número de ovos por desova, período de incubação, profundidade do ninho) serão apresentados com valores típicos, em torno dos quais a faixa possível de valores em populações naturais usualmente se situa. Os valores reais podem variar intra e inter-populações. Valores destes parâmetros para diferentes populações podem ser obtidos nas referências mencionadas acima.

2.1. Biologia reprodutiva

As tartarugas *Caretta caretta*, como as demais tartarugas marinhas, vivem inteiramente no mar, só estando em terra, na praia, em dois momentos do seu ciclo de vida: as fêmeas adultas, por ocasião da desova, e os filhotes recém-nascidos, quando saem do ninho e fazem sua breve corrida até o mar. Os machos, após o nascimento e a entrada no mar, não mais retornam à terra.

A vida de uma tartaruga *Caretta caretta* fêmea adulta reprodutivamente ativa pode ser vista como um ir e vir entre duas áreas do ambiente marinho separadas por uma distância que pode chegar a mais de 2000 km. Uma das áreas é a *área de alimentação*, onde ela reside a maior parte de tempo, e a outra é a *área de*

desova, contígua à praia onde a desova é realizada e onde permanece pelo período necessário às desovas. Uma *Caretta caretta* fêmea possui fidelidade a uma praia de desova, raramente desovando em uma praia diferente daquela em que o fez pela primeira vez. Estudos recentes, com a análise de DNA mitocondrial, indicam que esta fidelidade é uma fidelidade à praia de nascimento (Bowen *et al.*, 1993; Bowen *et al.*, 1994). Estudos utilizando a metodologia de marcação das tartarugas mostram que existe também uma fidelidade à área de alimentação (Limpus *et al.*, 1992).

As desovas das tartarugas *Caretta caretta* concentram-se em um período de alguns meses na época mais quente do ano. Este período de desova, dependente basicamente da latitude, é denominado de *temporada de desova*. As fêmeas desovam geralmente à noite, usualmente bem após o pôr do sol.

Uma tartaruga *Caretta caretta* pode desovar uma ou mais vezes em uma mesma temporada de desova. O número máximo de desovas observadas em uma mesma temporada para uma tartaruga desta espécie é de sete desovas, registrado nos Estados Unidos em 1979 (Lenarz *et al.*, 1981). O intervalo de tempo entre duas desovas consecutivas é denominado de *período internidal*, que para as tartarugas *Caretta caretta* é de aproximadamente 14 dias.

Raramente uma tartaruga *Caretta caretta* desova em anos consecutivos. O intervalo de tempo entre duas temporadas de desova consecutivas para uma mesma tartaruga é denominado de *intervalo de remigração*, que para as *Caretta caretta* é geralmente de 2 a 4 anos. Um fenômeno observado em diversas populações é que a maioria das fêmeas é vista apenas em uma única temporada de desova. Mesmo levando-se em conta a perda de marcas, a cobertura

incompleta do trabalho de marcação na praia, a possibilidade de desovas não registradas em praias adjacentes à praia sendo estudada e o nível de mortalidade admitido para as tartarugas marinhas, não há uma explicação aceita para a alta proporção de fêmeas que não é vista após sua primeira temporada de desova (Hughes, 1974b, 1982; Meylan, 1982; Richardson, 1982).

Uma desova é um processo composto por várias etapas. Uma *Caretta caretta*, antes de escavar o ninho, faz o que se denomina de *cama*, uma depressão na areia que tem a finalidade de acomodar o corpo da tartaruga durante a desova. O ninho propriamente dito, onde são depositados os ovos, é uma cavidade com cerca de 55 cm de profundidade, cavada com as nadadeiras posteriores. Ao terminar a desova, após cobrir o ninho, a tartaruga procura disfarçar o ponto exato de sua localização por meio de uma grande movimentação de areia em torno da área onde foi feita a cama, retornando a seguir ao mar. Todo o processo, desde a chegada à praia até a entrada de volta ao mar, dura cerca de 1 a 2 horas.

Nem sempre a saída da tartaruga do mar resulta na deposição de ovos. O processo de desova pode ser interrompido, por fatores diversos, em qualquer de suas etapas. Alguns dos fatores de interrupção são presença de luzes na praia, movimentação próxima de pessoas ou animais, encontro de grandes obstruções (pedras, troncos de árvores etc.) e perturbação direta da tartaruga, como por exemplo em atividades de marcação por pesquisadores na praia.

Alguns exemplos de situações de interrupção são: (1) Uma saída do mar com um retorno imediato à água, sem haver confecção da cama. O rastro de um movimento como este, em geral com a forma de um U, é conhecido como *meia-lua*. (2) A tartaruga pode confeccionar uma ou mais camas, em cada uma

sem haver escavação de ninho, mas havendo toda a movimentação de areia de uma desova completa, e retornar ao mar sem ter depositado ovos. (3) A tartaruga pode fazer a cama e o próprio ninho e retornar ao mar sem ter desovado. (4) A tartaruga pode iniciar a deposição dos ovos e interromper a desova no meio da deposição. Esta última possibilidade raramente acontece, pois é necessário que haja uma perturbação extremamente forte para que uma *Caretta caretta* interrompa o processo de desova uma vez que a deposição dos ovos tenha se iniciado. Caso uma *Caretta caretta* tenha vindo à praia e não realizado ou interrompido a desova, ela retornará à praia para nova tentativa, ou para completar a desova, em poucos dias, muito freqüentemente na mesma noite ou na noite seguinte à da tentativa frustrada.

A cada desova, um total em torno de 120 ovos são depositados no ninho escavado na areia, em geral em um ponto acima da linha da maré máxima. A fêmea retorna ao mar imediatamente após a desova, não tendo nenhum cuidado posterior com os ovos ou com os filhotes. Dos ovos depositados, em geral cerca de 55 a 80% resultam em filhotes vivos ao final de um intervalo de tempo, o período de incubação, que é da ordem de 55 dias. Nem sempre um ninho consegue chegar intacto ao final do período de incubação. Predadores (caranguejos, raposas etc., além da predação humana), inundação do ninho pelo mar e erosão da praia pela maré ou por tempestades são algumas das causas que podem fazer com que um ninho não chegue ao final da incubação ou, chegando, tenha uma produção de filhotes abaixo do que seria esperado em condições normais.

Uma tartaruga *Caretta caretta* fêmea adulta pesa em média de 100 a 120 kg, podendo chegar a 180 kg, e tem um comprimento médio da carapaça de cerca de 1 m. Estimativas para o tempo de maturação reprodutiva das *Caretta caretta* fêmeas, obtidas por meio de curvas de crescimento, chegam a mais de 35 anos (Foster, 1994; Frazer *et al.*, 1994). Estimativas para a longevidade das tartarugas *Caretta caretta* fêmeas, obtidas por técnicas de tabela de vida com base em dados de marcação colhidos em praias de reprodução nos Estados Unidos, indicam uma idade máxima da ordem de 62 anos (Frazer, 1983, 1986).

CAPÍTULO 3

Métodos computacionais e dados utilizados nas simulações

Neste capítulo serão apresentados: (1) os sistemas de computação e linguagens de programação utilizados neste trabalho, (2) algumas rotinas básicas empregadas nas simulações (geradores de números aleatórios e a rotina de otimização), (3) as distribuições de probabilidade utilizadas nas simulações como modelos para a distribuição do número de ninhos depositados por uma tartaruga em uma temporada e (4) a metodologia de construção de arquivos de dados para as simulações.

3.1. Computadores e linguagens de programação

Os programas que requereram o emprego de rotina de otimização (sempre que tenha havido estimação segundo o método do Capítulo 5), os programas para teste de quase-convexidade (Capítulo 5) e as simulações do Capítulo 10 foram realizados na linguagem AIX/Fortran 2.3 (International Business Machines Corporation, 1992) e executados no sistema paralelo IBM SP2 do Laboratório Nacional de Computação Científica (LNCC/CNPq), no Rio de Janeiro.

Os programas que não requereram o uso de rotina de otimização (simulações em que não houve estimação (Capítulo 4) e programas para análise de resultados)

foram realizados na linguagem Turbo Pascal 7.0 para DOS (Borland International Inc., 1992) e executados em computadores de tipo IBM-PC modelos 486 e Pentium.

O tratamento inicial aplicado aos dados do Projeto TAMAR (Capítulo 10), que estão em arquivos de tipo .DBF (padrão Dbase), foi feito por programas escritos na linguagem CA-Clipper 5.2 (Computer Associates International Inc., 1992), rodando em DOS em computadores de tipo IBM-PC modelos 486 e Pentium.

3.2. Rotinas-chave nas simulações

Geradores de números aleatórios. As simulações feitas em Turbo Pascal / DOS utilizaram como gerador de números aleatórios o gerador desenvolvido por L'Ecuyer (1988, Figura 3). Para as simulações realizadas em Fortran no sistema IBM SP2, o gerador de números aleatórios foi a rotina G05CAF da biblioteca de rotinas em Fortran (Mark 16) do NAG (Numerical Algorithms Group, 1993, volume 10). Em todas as simulações a semente inicial do gerador de números aleatórios foi uma função do relógio do computador.

Otimização. A minimização com restrições de uma função não-linear, necessária no processo de estimação por mínimos quadrados (Capítulo 5), foi feita através da rotina E04UPF da biblioteca de rotinas em Fortran (Mark 16) do NAG (Numerical Algorithms Group, 1993, volume 4).

3.3. Distribuições-modelo do número de ninhos por tartaruga

As simulações apresentadas nos Capítulos 4, 6, 7, 8, 9 e 10 utilizaram no total 32 modelos diferentes para a distribuição de probabilidade do número real

de ninhos depositados por uma tartaruga em uma temporada. Cada uma destas distribuições é definida por dois fatores: (1) um padrão (formato) básico para a distribuição, designado por um código formado por duas letras maiúsculas (ex: padrão RI) e (2) o número máximo de ninhos por tartaruga que é admitido pela distribuição, designado por ω_d e que pode variar conforme a distribuição (ex: $\omega_d = 6$). Uma *distribuição-modelo* é então designada por um código, com três caracteres, que representa a combinação padrão- ω_d que a define. Assim, como exemplo, a distribuição-modelo designada por RI6 segue o padrão RI e pressupõe um máximo de 6 ninhos por tartaruga em uma temporada.

As distribuições-modelo foram construídas com $\omega_d = 5, 6, 7$ e 8 , faixa de valores compatível com os números máximos de ninhos por tartaruga em uma temporada registrados na literatura para as tartarugas *Caretta caretta* (Dodd, 1988) e *Chelonia mydas* (Pritchard e Trebbau, 1984). Foram utilizados 8 padrões básicos na construção das distribuições-modelo, que estão apresentados na Figura 3.1 para $\omega_d = 7$. As distribuições com $\omega_d = 5, 6$ e 8 estão apresentadas no Apêndice A. Apenas o padrão RI é proveniente de dados reais. Os outros padrões são artificiais, escolhidos de forma arbitrária e criados com o intuito de testar as propriedades do modelo para estimação de população proposto no Capítulo 5 e para servirem como modelos para as simulações apresentadas no Capítulo 4. Os padrões representam diferentes condições de assimetria e unimodalidade ou multimodalidade possíveis de serem encontrados na distribuição do número de ninhos por tartaruga.

A distribuição RI7, publicada por Richardson (1982, Figura 9), representa o número suposto de ninhos por tartaruga para as tartarugas marinhas *Caretta caretta* desovando em Little Cumberland Island, Georgia, USA, e foi elaborada

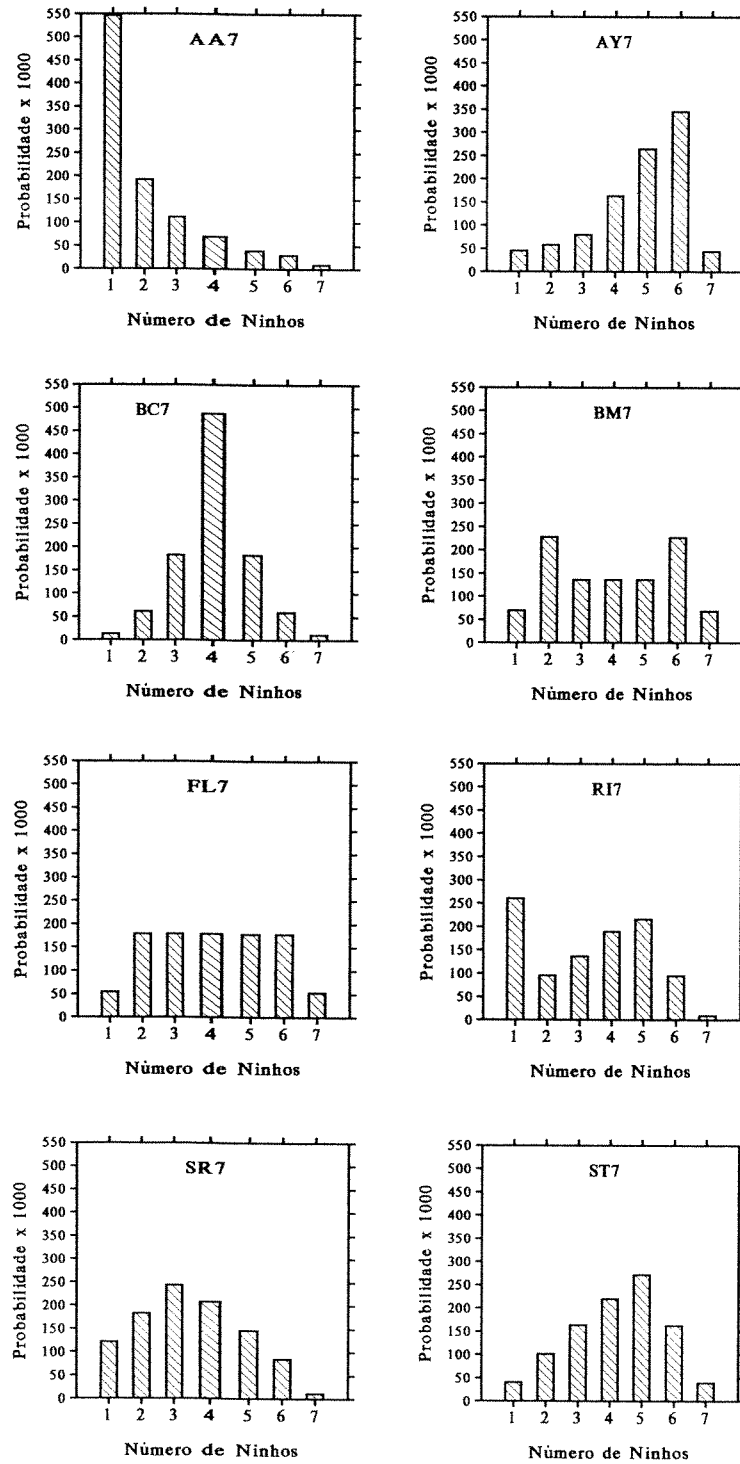


Figura 3.1. Distribuições-modelo para o número de ninhos por tartaruga com $\omega_d = 7$: AA7, AY7, BC7, BM7, FL7, RI7, SR7 e ST7.

com dados obtidos no período 1977-1981. O número suposto de ninhos para uma tartaruga é o número observado de ninhos acrescido dos ninhos supostamente colocados pela tartaruga e não observados. Estes ninhos supostamente não observados são detectados pela análise do intervalo de tempo ocorrido entre duas desovas observadas consecutivas. Em Little Cumberland Island, em 1979, foi registrado o número máximo de desovas por temporada já observado para uma tartaruga *Caretta caretta*, que é de 7 ninhos (Lenarz *et al.*, 1981). Devido à alta cobertura do trabalho de campo em Little Cumberland Island, sempre superior a 95% entre 1977 e 1981 (Richardson, 1982), esta distribuição RI7, uma distribuição bimodal, é muito próxima da distribuição real de ninhos por tartaruga naquela praia. Para esta distribuição, o número médio de ninhos por tartaruga é 3.32.

Para $\omega_d = 5, 6$ e 8 , as distribuições seguem o padrão (formato) da distribuição correspondente com $\omega_d = 7$ (veja o Apêndice A). Assim, as distribuições RI5, RI6 e RI8, conforme o padrão da distribuição RI7, são distribuições bimodais, as distribuições AA5, AA6 e AA8, da mesma forma que a distribuição AA7, são distribuições altamente assimétricas com $p(k) > p(k+1)$, $k = 1, 2, \dots, \omega_d - 1$, ($p(k)$ = probabilidade de uma tartaruga desovar k vezes na temporada) e assim por diante em relação aos outros padrões. A Figura 3.2 mostra, como exemplo, as distribuições RI5, RI6 e RI8 lado a lado com a distribuição RI7.

3.4. Construção dos arquivos de dados usados nas simulações

As simulações realizadas neste trabalho foram feitas a partir de populações finitas de tartarugas, com T , o número de tartarugas em cada população,

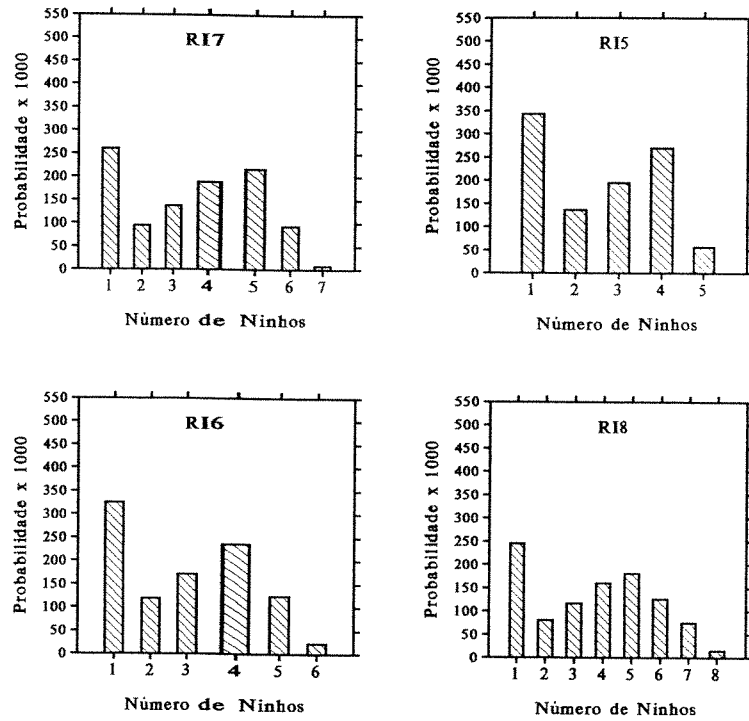


Figura 3.2. Distribuições-modelo seguindo o padrão RI com diferentes valores para ω_d : RI7 (acima, à esquerda, a mesma da Figura 3.1), RI5, RI6 e RI8.

pré-fixado. Para cada combinação dos padrões básicos AA, AY, BC, BM, FL, RI, SR e ST com os diferentes valores de ω_d , $\omega_d = 5, 6, 7$ e 8 , foram criados cinco diferentes arquivos de dados, cada arquivo representando uma população de tartarugas com o valor de T pré-fixado (uma *população-modelo*) e tendo o número de ninhos por tartaruga obedecendo aproximadamente à distribuição-modelo considerada. Os cinco arquivos têm os seguintes valores para T : 10, 20, 40, 100 e 250. No total, foram criadas então 160 populações-modelo diferentes ($8 \text{ padrões} \times 4 \text{ valores para } \omega_d \times 5 \text{ valores para } T$). Cada população-modelo é representada por uma sequência $\{X_j\}_{j=1}^T$, onde X_j é o número de ninhos depositados pela tartaruga j na temporada, $X_j \in \{1, 2, \dots, \omega_d\}$.

Como as populações-modelo criadas são finitas, não seria em geral possível construir uma população com a distribuição do número de ninhos por tartaruga nesta população sendo igual à distribuição-modelo. O algoritmo de construção da população finita buscou, entre todas as distribuições geradas por T tartarugas, aquela que mais se aproximasse da distribuição-modelo no sentido da norma euclideana no espaço R^{ω_d} , considerando-se uma distribuição como um ponto no espaço R^{ω_d} , o espaço vetorial dos pontos com coordenadas reais com dimensão ω_d , sendo $p(k)$, a probabilidade de haver k ninhos na temporada, a coordenada k deste ponto em R^{ω_d} . Assim, um arquivo de dados, representando uma população-modelo, não foi obtido por amostragem aleatória da sua distribuição-modelo, mas tem a população seguindo o mais fielmente possível a distribuição-modelo. Como exemplo de arquivo de dados, as 40 tartarugas que compõem a população com $T = 40$ seguindo a distribuição-modelo RI7 são representadas pela sequência $\{X_j\}_{j=1}^{40}$ abaixo:

1	1	1	1	1	1	1	1	1	1	2	2	2	2
3	3	3	3	3	4	4	4	4	4	4	4	4	
5	5	5	5	5	5	5	5	5	5	6	6	6	6

A Figura 3.3 mostra lado a lado a distribuição-modelo RI7 com a distribuição do número de ninhos por tartaruga para esta população de 40 tartarugas apresentada logo acima.

Se T , o tamanho da população-modelo construída para o arquivo de dados, é relativamente pequeno, pode ocorrer que o número máximo de ninhos para uma tartaruga na população-modelo, denotado por ω_p , seja menor que o ω_d da distribuição-modelo, caso $p(\omega_d)$ seja suficientemente pequeno. Ou seja, pode

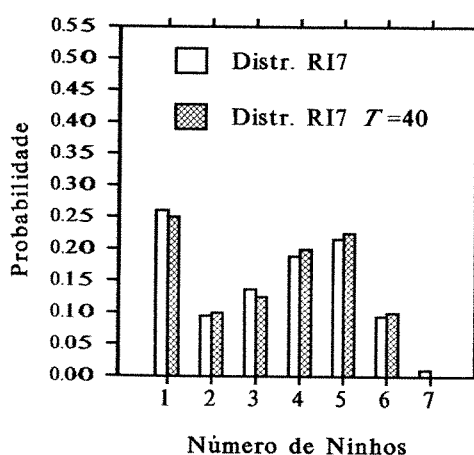


Figura 3.3. Comparação entre a distribuição de probabilidade RI7 e a distribuição do número de ninhos por tartaruga na população-modelo com $T = 40$ seguindo esta distribuição RI7.

ocorrer que nenhuma tartaruga na população-modelo construída alcance o número máximo de ninhos admitido pela distribuição-modelo. No exemplo apresentado logo acima, ω_p é igual a 6, portanto menor que o ω_d da distribuição RI7, que é igual a 7. Para a distribuição RI7, $p(7) = 0.009$, probabilidade muito pequena para que uma população com $T = 40$ contivesse uma tartaruga depositando 7 ninhos na temporada.

CAPÍTULO 4

Uma análise do problema

O estimador classicamente utilizado na estimação do número de tartarugas marinhas desovando em uma praia em uma temporada de desova (Schulz, 1975; Le Gall *et al.*, 1986; Mortimer e Carr, 1987; Mortimer 1988) é dado pela equação

$$\hat{T} = \hat{N} / \hat{\mu}, \quad (4.1)$$

onde $\hat{T}$ é o número estimado de fêmeas desovando na temporada, $\hat{N}$ é o número total estimado de ninhos colocados na praia na temporada pelo conjunto das tartarugas e $\hat{\mu}$ é o número médio estimado de ninhos depositados por fêmea na temporada. Se o número total de ninhos na temporada, N , é conhecido, não necessitando ser estimado, o problema de estimar o número de tartarugas desovando na temporada é equivalente ao problema de estimar μ , o número médio de ninhos depositados por fêmea na temporada.

No presente capítulo será feita, por meio de exemplos obtidos por simulações em computador, uma discussão sobre a natureza do problema da estimação do número médio real de ninhos por tartaruga, ou seja, sobre o problema de estimação referente ao denominador do lado direito da equação (4.1).

4.1. O processo de observação na praia

Se uma tartaruga marinha coloca um determinado número de ninhos na praia na temporada, quanto menor a probabilidade de que esta tartaruga seja observada no momento de uma desova (isto é, seja flagrada na praia pela equipe de campo no momento da desova, o que permite associar o ninho à tartaruga), menor será provavelmente o número de ninhos reconhecidos como pertencentes a esta tartaruga na temporada. A probabilidade de que a tartaruga seja observada no momento de uma desova é uma função direta do esforço de trabalho de campo (medido por exemplo pelo número diário de horas de patrulhamento da praia). Se este esforço de trabalho de campo for relativamente baixo, será alta a probabilidade de que cada ninho desta tartaruga não tenha a tartaruga-mãe identificada.

Vamos denominar de *tartaruga observada* uma tartaruga que foi encontrada ao menos uma vez na temporada pela equipe de campo no momento de uma de suas desovas, e de *ninho observado* um ninho do qual se conhece a tartaruga-mãe, ou seja, a tartaruga-mãe foi encontrada pela equipe de campo quando depositava este ninho, podendo-se assim associar o ninho à tartaruga.

Seja p definida como a probabilidade de encontro de um ninho qualquer no momento em que a desova é realizada, válida para o conjunto dos ninhos da temporada. É de se esperar que o número médio de ninhos observados por tartaruga, entre as tartarugas observadas, seja menor do que a média real de ninhos por tartaruga, se p for menor do que 1.0. Este fato já foi reconhecido por diversos pesquisadores (Pritchard e Trebbau, 1984; Frazer e Richardson, 1985; Mortimer e Carr, 1987).

No entanto, pode-se ter, adicionalmente, uma idéia qualitativa do comportamento do número médio de ninhos observados entre as tartarugas observadas como função de p : (1) quanto menor a probabilidade p , ou seja, quanto menor o esforço de trabalho de campo à noite, pode-se esperar que, para as tartarugas que tiveram alguma de suas desovas observadas, menor será a média do número de ninhos observados por tartaruga, ou seja, o número médio de ninhos observados por tartaruga observada deve ser uma função decrescente de p ; (2) no limite, quando $p \rightarrow 0$, pode-se esperar que o número médio de ninhos observados por tartaruga observada tenderá a 1.0, qualquer que seja o número médio real de ninhos por tartaruga na população, pois será muito pouco provável que uma mesma fêmea seja encontrada mais de uma vez na temporada caso o esforço de trabalho de campo seja extremamente pequeno. Será mostrado a seguir, por meio de simulações em computador, como o número médio de ninhos observados entre as tartarugas observadas depende de forma geral da probabilidade p .

4.2. Metodologia de simulação

O objetivo principal das simulações é obter, para diferentes populações-modelo de tartarugas, como função de p , $\bar{\mu}_o(p)$, o número médio esperado de ninhos observados para as tartarugas observadas, com p variando entre 1.0 e um valor próximo de 0.0. As simulações foram feitas utilizando-se populações-modelo construídas segundo metodologia descrita no Capítulo 3, utilizando-se as distribuições AA7, AY7, BC7, BM7, FL7, RI7, SR7 e ST7 com $T = 40$ e diversos valores para p . Para cada população-modelo, para cada valor de p , foram realizadas 2000 simulações de observações em uma temporada. Cada uma destas

simulações (onde a população-modelo é assumida como real) representa uma temporada completa de observações feitas pela equipe de campo na praia, condicionada a que tenha havido pelo menos uma tartaruga observada na temporada, sendo que a probabilidade de observação de cada ninho é constante na temporada. Para cada população-modelo, as simulações foram feitas de acordo com o seguinte algoritmo:

1. Do arquivo de dados relativo à população-modelo, conhece-se T , o número de tartarugas na população, e $\{X_j\}_{j=1}^T$, onde X_j é o número de ninhos colocados na temporada pela tartaruga j , $j = 1, \dots, T$.

2. Para cada valor de p , $p = 0.98, 0.95, 0.90, 0.70, 0.50, 0.30, 0.20, 0.15, 0.10, 0.05, 0.03$ e 0.01 , repetir os itens 2.1 - 2.2 até que haja 2000 valores $\mu_o(p)$ calculados no item 2.2 ($\mu_o(p)$ é uma variável aleatória).

2.1. Para $j = 1, \dots, T$, gerar $Y_j \sim \text{Bin}(X_j, p)$, variável aleatória seguindo distribuição binomial com parâmetros X_j e p , onde Y_j representa o número de ninhos observados da tartaruga j .

2.2. Caso não exista nenhum $Y_j > 0$, $j = 1, \dots, T$, retornar ao item 2.1. Se

existir algum $Y_j > 0$, calcular $\mu_o(p) = \sum_{j=1}^T Y_j / \sum_{j=1}^T I_{Y_j > 0}$, onde

$$I_{Y_j > 0} = \begin{cases} 1 & \text{se } Y_j > 0 \\ 0 & \text{se } Y_j = 0 \end{cases}. \quad \text{Assim, } \mu_o(p) \text{ é a média de ninhos observados}$$

por tartaruga observada nesta temporada simulada.

3. Da distribuição dos 2000 $\mu_o(p)$ calculados no item 2 para cada valor de p , calcular a média, $\bar{\mu}_o(p)$, e os percentis 2.5 e 97.5 para este p .

Para $p = 1.0$, $\bar{\mu}_o(1.0)$ e os respectivos percentis não foram calculados por simulação. Para $p = 1.0$, $\mu_o(1.0) = \mu$ necessariamente, onde μ é a média de

ninhos por tartaruga na população-modelo, de modo que, para este valor de p , $\bar{\mu}_o(1.0) = \mu$ e os percentis da distribuição dos 2000 $\mu_o(1.0)$ são também iguais a μ .

4.3. O processo de observação na praia: resultados

A Figura 4.1 mostra o número médio esperado de ninhos observados por tartaruga observada, $\bar{\mu}_o(p)$, e também os percentis 2.5 e 97.5 da distribuição dos $\mu_o(p)$, para a simulação feita com a distribuição RI7 e $T = 40$. As simulações realizadas com as outras distribuições-modelo produziram resultados análogos aos da Figura 4.1. A Figura 4.2 mostra $\bar{\mu}_o(p)$ (não são apresentados os percentis) para as oito distribuições-modelo AA7, AY7, BC7, BM7, FL7, RI7, SR7 e ST7 com $T = 40$ (a curva para a distribuição RI7, a terceira curva de baixo para cima, é a mesma curva apresentada na Figura 4.1. Observar que as escalas verticais das Figuras 4.1 e 4.2 são diferentes).

Na Figura 4.2, o quadrado preto representa o valor $\mu_o = 1.89$ obtido em uma população hipotética de tartarugas marinhas em uma temporada onde $p = 0.34$. Este ponto com coordenadas $(0.34, 1.89)$ foi incluído na figura como exemplo de como estaria representada no gráfico $p \times \mu$ uma temporada de observações realizada em uma população real de tartarugas (veja a discussão a seguir).

4.4. O processo de observação na praia: discussão

As Figuras 4.1 e 4.2 permitem uma visualização da dependência do número médio de ninhos observados por tartaruga observada em relação a p . A Figura 4.1 mostra que, como previsto, em geral $\mu_o < \mu$ se $p < 1.0$, onde μ_o é o número

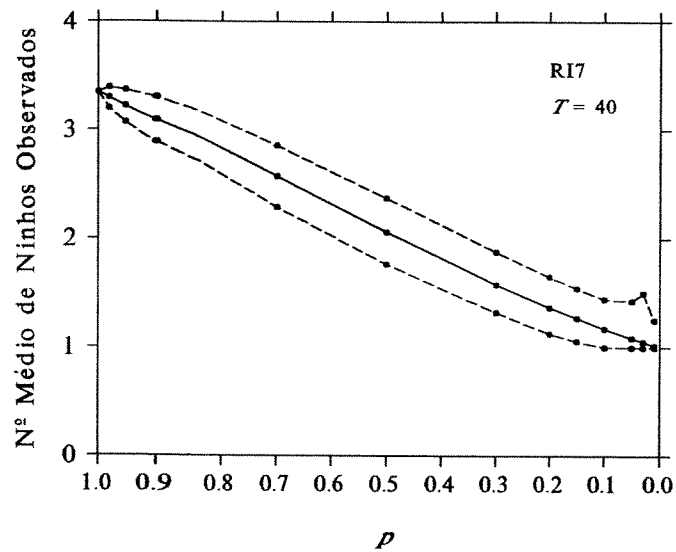


Figura 4.1. Número médio esperado de ninhos observados por tartaruga observada em função de p (linha cheia), resultados obtidos por simulação, condicionados a que tenha havido ao menos uma tartaruga observada na temporada. Distribuição RI7, $T = 40$. São mostrados também os percentis 2.5 e 97.5 (linhas tracejadas) da distribuição do número médio de ninhos observados para cada valor de p . Observar a escala reversa no eixo horizontal.

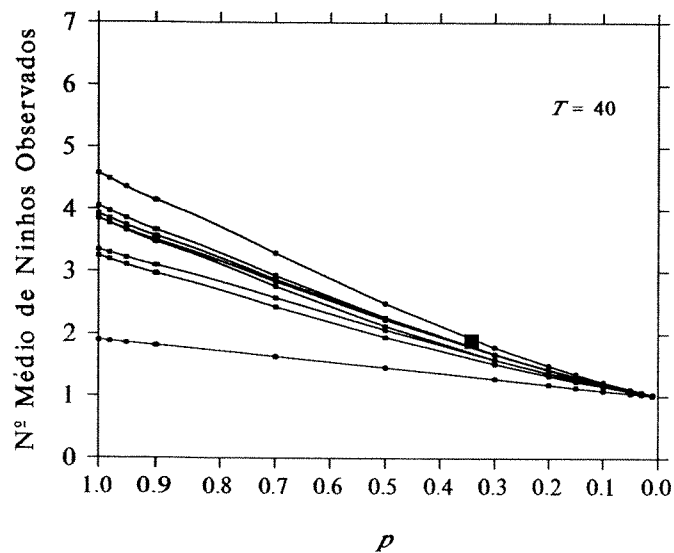


Figura 4.2. Número médio esperado de ninhos observados por tartaruga observada em função de p , resultados obtidos por simulação, condicionados a que tenha havido ao menos uma tartaruga observada na temporada. As linhas representam $\bar{\mu}_o(p)$ para cada uma das oito distribuições (de cima para baixo) AY7, ST7, FL7, BC7, BM7, RI7, SR7 e AA7 com $T = 40$. Observar a escala reversa no eixo horizontal. O quadrado preto representa uma observação hipotética com $p = 0.34$ e $\mu_o = 1.89$ (veja o texto).

médio de ninhos observados entre as tartarugas observadas, o que faz com que, de forma geral, utilizar μ_o no denominador do lado direito da equação (4.1) equivale, caso N seja conhecido, a superestimar o número de tartarugas desovando na praia na temporada.

As Figuras 4.1 e 4.2 mostram que o problema de estimar o número médio real de ninhos por tartaruga pode ser entendido como um problema inverso: dado um ponto (p, μ_o) no gráfico $p \times \mu$ com $p < 1.0$ (por exemplo, o quadrado preto com coordenadas (0.34, 1.89) na Figura 4.2), o problema é estimar um ponto da forma $(1.0, \mu)$ que seja correspondente (ou seja, relacionado pelo processo de observação na praia) ao ponto (p, μ_o) dado inicialmente. A Figura 4.2 sugere que quanto mais próximo de $p = 0.0$ estiver o p do ponto dado inicialmente (o ponto obtido pela observação na praia) mais difícil será fazer a estimativa, ou seja, maior será o erro associado à estimativa. Em outros termos, quanto mais próximo de $p = 0.0$ estiver o p do ponto dado inicialmente, maior será a incerteza quanto à coordenada μ do ponto $(1.0, \mu)$ obtido pelo processo de estimativa.

4.5. A transformação da distribuição de probabilidade

A Figura 4.2 mostra que a média do número de ninhos observados entre as tartarugas observadas tende a 1.0 quando p tende a 0.0. Ocorre que uma média é na realidade a média de uma distribuição de probabilidade. Como o número de ninhos que uma tartaruga pode depositar na temporada não pode ser menor do que 1, a convergência da média do número de ninhos observados entre as tartarugas observadas para 1.0 quando p tende a 0.0 significa que a distribuição de probabilidade do número de ninhos observados entre as tartarugas observadas tende, quando p tende a 0.0, à distribuição de probabilidade que tem todo o peso

em 1 ninho, ou seja, à distribuição de probabilidade tal que $p(1) = 1.0$ e $p(k) = 0.0$ se $k > 1$. Em outros termos, se as distribuições de probabilidade do número de ninhos por tartaruga são representadas graficamente, a exemplo da Figura 3.1, o fato da média das distribuições tender a 1.0 quando p tende a 0.0 significa que, no gráfico, a distribuição do número de ninhos observados entre as tartarugas observadas desloca-se para a esquerda quando p tende a 0.0.

Este deslocamento para a esquerda da distribuição do número de ninhos observados entre as tartarugas observadas quando p tende a 0.0 pode também ser visualizado por meio de simulações em computador. As simulações descritas na Seção 4.2 permitem, adicionalmente ao item 2.2 do algoritmo daquela Seção, que se calcule a distribuição de probabilidade do número de ninhos observados entre as tartarugas observadas, na situação em que exista algum $Y_j > 0$. As probabilidades serão calculadas por $p(k) = \#\{Y_j \mid Y_j = k\} / \sum_{j=1}^T I_{Y_j > 0}$, $k = 1, \dots, \omega_p$, onde ω_p é o número máximo de ninhos por tartaruga na população-modelo.

A Figura 4.3 mostra uma realização da simulação de observações em uma temporada (ou seja, uma única passagem completa pelo item 2 do algoritmo da Seção 4.2) para cada um dos seguintes valores de p : 1.00, 0.90, 0.50 e 0.20, para a população-modelo seguindo a distribuição RI7 com $T = 40$ (é a população representada pela sequência $\{X_j\}_{j=1}^{40}$ apresentada no Capítulo 3, página 21, e mostrada na Figura 3.3). Para cada uma destas realizações é apresentada a distribuição do número de ninhos observados entre as tartarugas observadas. A distribuição de probabilidade do número de ninhos por tartaruga na população-modelo está representada na Figura 4.3 pelo gráfico correspondente a $p = 1.0$, valor de p que ocorre quando todas as tartarugas da população são

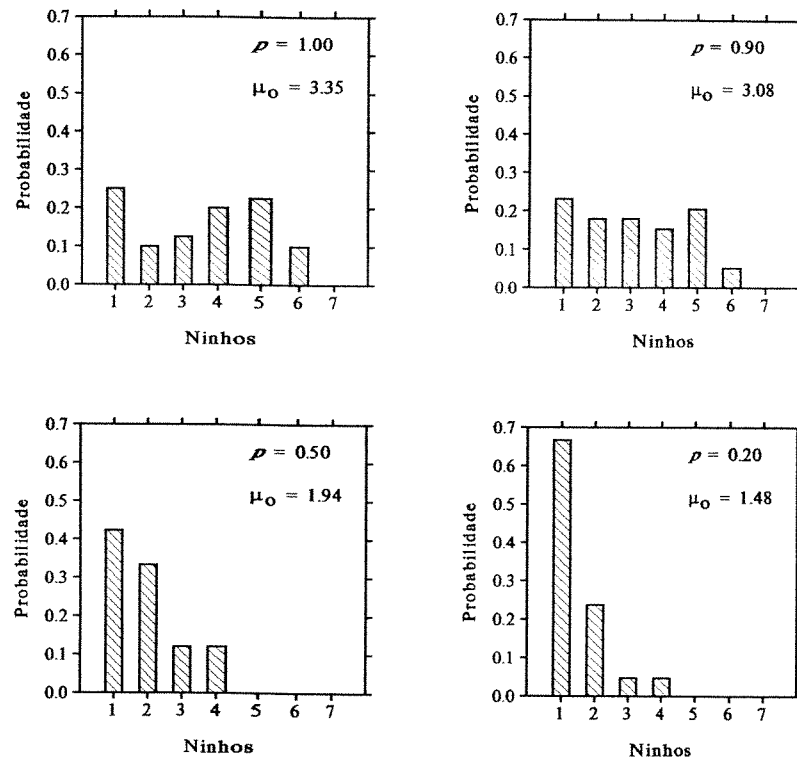


Figura 4.3. Distribuições de probabilidade do número de ninhos observados entre as tartarugas observadas, população-modelo seguindo a distribuição RI7, $T = 40$, resultados obtidos por simulação. Cada simulação é uma realização de uma temporada completa de observações na praia. Em cada gráfico estão mostrados o valor de p usado na simulação e o valor de μ_o , a média do número de ninhos observados entre as tartarugas observadas na simulação. O gráfico para $p = 1.0$ mostra a distribuição de probabilidade no número de ninhos por tartaruga na própria população-modelo.

observadas pela equipe de campo em todas as suas desovas na temporada. Na Figura 4.3 são apresentados também os valores de μ_o referentes a cada uma das distribuições apresentadas na figura.

A análise da transformação sofrida pela distribuição de probabilidade do número de ninhos por tartaruga quando $p < 1.0$, ilustrada pela Figura 4.3, será a base do modelo apresentado no próximo capítulo para a estimação do número de tartarugas desovando em uma praia em uma temporada.

CAPÍTULO 5

O modelo

Conforme discutido no final do Capítulo 4, o processo de observação pode ser compreendido como uma transformação ocorrendo entre a distribuição de probabilidade do número de ninhos por tartaruga na população, cuja média é μ , e a distribuição do número de ninhos observados entre as tartarugas observadas, cuja média é μ_o . Neste capítulo será proposto um modelo para a estimação da média μ da população feita a partir do conhecimento da distribuição do número de ninhos observados entre as tartarugas observadas e do número de ninhos depositados na temporada pelo conjunto das tartarugas da população. O modelo tem como base a análise, realizada com o uso da teoria de probabilidades, da transformação entre distribuições mencionada acima e é aplicável à situação em que N , o número total de ninhos depositados na praia na temporada, é conhecido exatamente. Dado que N é conhecido, a estimativa para o número de tartarugas desovando na temporada, $\hat{T}$, é obtida imediatamente da estimativa $\hat{\mu}$ para o número médio de ninhos produzida pelo modelo. No Capítulo 11 será proposta uma extensão do modelo à situação em que N também é estimado.

5.1. Definição do estimador

O estimador clássico para o número de tartarugas desovando em uma

temporada (equação (4.1)), na situação em que N é conhecido, toma a forma

$$\hat{T} = N/\hat{\mu}, \quad (5.1)$$

onde $\hat{T}$ é o número total estimado de tartarugas, N é o número total de ninhos e $\hat{\mu}$ é o número médio estimado de ninhos por tartaruga, todos os valores sendo referentes a uma mesma temporada. Como N é conhecido em (5.1), estimar T é equivalente a estimar μ . Na construção apresentada a seguir, N será suposto conhecido exatamente, ou seja, com erro igual a zero.

Vamos denotar por ω o número máximo de ninhos que uma tartaruga pode depositar em uma temporada. Este valor ω será suposto conhecido na construção feita a seguir, não sendo estimado pelo modelo. Será feita posteriormente (Capítulo 7) uma análise de sensibilidade do modelo em relação a esta determinação *a priori* de ω .

Vamos supor ainda que, através do processo de observação das tartarugas na praia, feito pela equipe de campo, cada ninho tenha uma probabilidade $p > 0$ de ser observado (atribuído a uma tartaruga-mãe). Esta probabilidade p de observação de cada ninho será estimada por $\hat{p}$, definido por

$$\hat{p} = N_o/N, \quad (5.2)$$

onde N_o é o número de ninhos observados na temporada.

Vamos denotar por F a distribuição do número de ninhos por tartaruga na população. Vista como elemento de R^ω , o espaço vetorial dos pontos com coordenadas reais com dimensão ω , a distribuição F equivale a um vetor

$$F = (\alpha_1, \alpha_2, \dots, \alpha_\omega), \quad (5.3)$$

onde α_k é a probabilidade de que uma tartaruga coloque k ninhos na temporada, $k = 1, \dots, \omega$.

Vamos denotar a distribuição observada por H_o . Esta distribuição equivale a um vetor

$$H_o = (\gamma_1^o, \gamma_2^o, \dots, \gamma_\omega^o), \quad (5.4)$$

onde γ_k^o é a probabilidade de que uma tartaruga observada tenha k ninhos observados, $k = 1, \dots, \omega$. As probabilidades γ_k^o são computadas a partir dos dados de campo por

$$\gamma_k^o = \frac{\text{número de tartarugas observadas com } k \text{ ninhos observados}}{\text{número total de tartarugas observadas}} \quad (5.5)$$

Um argumento baseado na teoria de probabilidades (Apêndice 5.A, neste capítulo) mostra que a distribuição $F = (\alpha_1, \dots, \alpha_\omega)$ transforma-se, se $p > 0$, na distribuição H definida por

$$H = (\gamma_1, \gamma_2, \dots, \gamma_\omega), \quad (5.6)$$

onde

$$\gamma_k = \sum_{i=1}^{\omega} b_{ki} \alpha_i / \left(1 - \sum_{i=1}^{\omega} b_{0i} \alpha_i \right), \quad k = 1, \dots, \omega, \quad (5.7)$$

sendo os coeficientes b_{ki} definidos por

$$b_{ki} = \begin{cases} \binom{i}{k} p^k (1-p)^{i-k} & \text{se } 0 \leq k \leq i \\ 0 & \text{se } k > i \end{cases}, \quad i = 1, \dots, \omega, \quad k = 0, 1, \dots, \omega. \quad (5.8)$$

Os γ_k da equação (5.7) definem uma função

$$f: \alpha\text{-espaço} \subseteq R^\omega \rightarrow \gamma\text{-espaço} \subseteq R^\omega \quad (5.9)$$

entre o conjunto das distribuições $F = (\alpha_1, \dots, \alpha_\omega)$, denominado de α -espaço, e o conjunto das distribuições $H = (\gamma_1, \dots, \gamma_\omega)$, denominado de γ -espaço. Esta função f (Figura 5.1) é dada por

$$f(\alpha_1, \dots, \alpha_\omega) = (\gamma_1, \dots, \gamma_\omega), \quad (5.10)$$

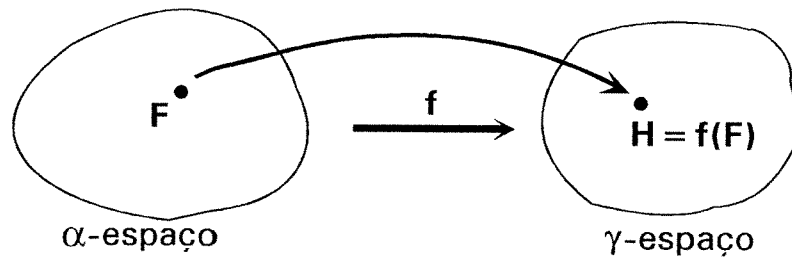


Figura 5.1.

onde os γ_k são definidos pela equação (5.7) e o α -espaço e o γ -espaço são simplexes em R^ω , definidos por

$$\begin{aligned} \alpha\text{-espaço} &= \gamma\text{-espaço} = \\ \{(x_1, \dots, x_\omega) \in R^\omega \mid 0 \leq x_k \leq 1, k=1, \dots, \omega \text{ e } x_1 + \dots + x_\omega = 1\}. \end{aligned} \quad (5.11)$$

A função f é um mapeamento do espaço das distribuições $F = \{\alpha_k\}_{k=1}^\omega$ (o α -espaço) no espaço das distribuições $H = \{\gamma_k\}_{k=1}^\omega$ (o γ -espaço). No caso de uma população de tartarugas T infinita, se $p > 0$, a função f mostraria como a distribuição F real se transformaria na distribuição H_o observada: H_o seria a imagem de F por f , $H_o = f(F)$. No entanto, para uma população T finita (que é o que existe no mundo real), se $p > 0$, a distribuição H_o observada não será a imagem $f(F)$ da distribuição F pela função f , mas sim uma aproximação desta imagem $f(F) = H$. Para uma distribuição F fixa no α -espaço existirá no γ -espaço uma distribuição amostral das distribuições observáveis para este $p > 0$ (Figura 5.2). A distribuição H_o observada é um ponto pertencente a esta distribuição amostral, é uma realização de um processo estocástico (a observação das tartarugas na praia) que, a partir da distribuição F real, gera a distribuição observada.

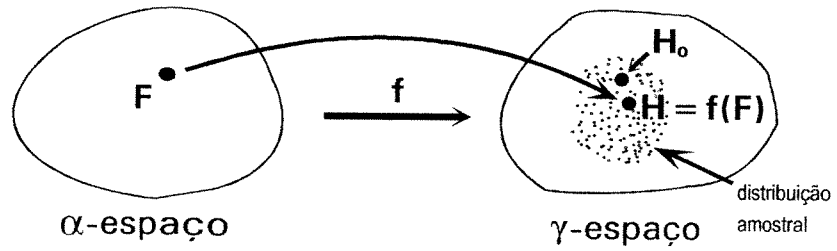


Figura 5.2.

Conhecida a distribuição H_o , a estimativa $\hat{F}$ da distribuição F será definida como aquela distribuição $\hat{F}$ no α -espaço tal que, entre todas as distribuições $f(F)$ no γ -espaço, onde $F \in \alpha$ -espaço, $f(\hat{F})$ seja a melhor aproximação possível (no sentido de menor distância) para H_o no γ -espaço, sendo a distância entre duas distribuições no γ -espaço medida pela norma euclidiana em R^{ω} (Figura 5.3). Formalmente, $\hat{F}$ é tal que

$$\|f(\hat{F}) - H_o\| = \min_{F \in \alpha\text{-espaço}} \|f(F) - H_o\|, \quad (5.12)$$

onde $\|\cdot\|$ é a norma euclidiana em R^{ω} . Em (5.12), no cálculo da função f ,

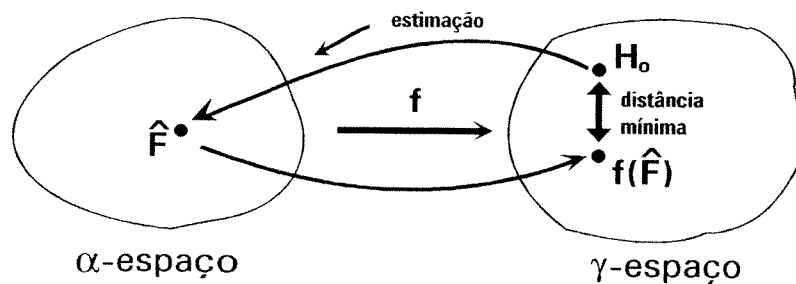


Figura 5.3.

os coeficientes b_{ki} da equação (5.7) são calculados utilizando-se, na equação (5.8), $\hat{p}$ definido pela equação (5.2), no lugar de p , que não é conhecido. (Na realidade, está se utilizando em (5.12), no lugar de f , uma função estimada $\hat{f}$, definida de modo análogo à definição de f por meio das equações (5.7) - (5.10), com $\hat{p}$ no lugar de p .)

Uma vez calculada a distribuição $\hat{F}$ por (5.12), a estimativa $\hat{\mu}$ para a média de ninhos por tartaruga na população será então definida por

$$\hat{\mu} = \sum_{k=1}^{\omega} k \hat{\alpha}_k, \quad (5.13)$$

onde os $\hat{\alpha}_k$ são as coordenadas da distribuição $\hat{F}$ enquanto elemento de R^{ω} . Uma vez calculado $\hat{\mu}$ por (5.13), $\hat{T}$ pode então ser calculado por (5.1), arredondado para o número inteiro mais próximo, dado que admite-se conhecido o valor de N .

5.2. Existência e unicidade da solução

A solução do problema de otimização definido em (5.12) existe e é única. Consideremos, com $p > 0$ e $H_o \in \gamma$ -espaço, a função $g : \alpha$ -espaço $\subseteq R^{\omega} \rightarrow R$ definida por $g(F) = \|f(F) - H_o\|$, $F \in \alpha$ -espaço, onde a função f é definida por (5.7) - (5.10).

Existência. O mínimo existe, pois a função g é contínua e definida em um subconjunto compacto de R^{ω} .

Unicidade. A função g tem a propriedade de ser estritamente quase-convexa, ou seja, se $F_1, F_2 \in \alpha$ -espaço, $F_1 \neq F_2$, e λ é um número real, $0 < \lambda < 1$, então $g(\lambda F_1 + (1 - \lambda)F_2) < \max\{g(F_1), g(F_2)\}$ (Avriel *et al.*, 1981). (Esta propriedade é também às vezes denominada na literatura por fortemente quase-convexa, ex. Bazaraa *et al.* (1993).) Pode ser demonstrado que, para uma

função estritamente quase-convexa definida em um conjunto convexo (o α -espaço é convexo), se um mínimo local existe (no caso da função g o mínimo existe), este mínimo é o mínimo global único no domínio da função (Diewert *et al.*, 1981, teorema 8).

A quase-convexidade da função g foi verificada em computador, para $2 \leq \omega \leq 12$, pelo exame direto da definição de quase-convexidade, de acordo com o seguinte algoritmo:

1. Para cada ω , $2 \leq \omega \leq 12$, executar o item 2.
2. Para cada um dos seguintes valores de p : 0.01 (0.01) 0.09, 0.10 (0.02) 0.20, 0.225 (0.025) 0.775, 0.80 (0.02) 0.90 e 0.91 (0.01) 0.99, executar o item 3.
3. Escolher aleatoriamente 2300 distribuições H_o no γ -espaço $\subseteq R^\omega$ e, para cada H_o escolhido, executar 2300 vezes o item 4. (Não há nenhuma razão especial para a escolha do número 2300.)
4. Fixados ω (item 1), p (item 2) e H_o (item 3), escolher aleatoriamente um par de pontos F_1 e F_2 no α -espaço $\subseteq R^\omega$, $F_1 \neq F_2$, e escolher aleatoriamente um número real λ , $0 < \lambda < 1$, e verificar então se a desigualdade $g(\lambda F_1 + (1 - \lambda)F_2) < \max\{g(F_1), g(F_2)\}$ é satisfeita para o conjunto $\{\omega, p, H_o, F_1, F_2, \lambda\}$.

No total, para $2 \leq \omega \leq 12$, a desigualdade foi verificada, e sempre satisfeita, cerca de 3.08 bilhões de vezes, em programas que consumiram aproximadamente 209 horas de tempo de processador no sistema IBM SP2.

5.3. Discussão

O estimador apresentado em (5.12) é definido por um processo de distância

mínima (Parr, 1985), onde a distribuição $f(F)$, uma aproximação para H_o no γ -espaço, é parametrizada por $\{\alpha_k\}_{k=1}^{\omega}$, com as restrições $0 \leq \alpha_k \leq 1$ e $\sum_{k=1}^{\omega} \alpha_k = 1$. Como esta parametrização cobre todas as distribuições existentes no α -espaço, nenhuma suposição é feita *a priori* sobre a forma da distribuição $F \in \alpha$ -espaço.

Em termos computacionais, estimar F por (5.12) significa um problema de minimização de uma função não-linear em um espaço real com ω dimensões com a restrição $F \in \alpha$ -espaço, onde o α -espaço é definido por (5.11).

Apêndice 5.A. Dedução da equação (5.7)

Suponhamos que T tartarugas desovaram na praia na temporada. Seja X_j o número de ninhos produzidos pela tartaruga j , $j = 1, \dots, T$, sendo $X_j \in \{1, 2, \dots, \omega\}$. Seja $F = \{\alpha_k\}_{k=1}^{\omega}$ a distribuição de probabilidade dos $\{X_j\}_{j=1}^T$,

$$\Pr(X_j = k) = \alpha_k, \quad \alpha_k \geq 0, \quad \sum_{k=1}^{\omega} \alpha_k = 1, \quad (5A.1)$$

onde

$$\alpha_k = \#\{X_j | X_j = k\} / T, \quad (5A.2)$$

ou seja, α_k é a proporção das tartarugas que colocaram k ninhos na temporada, $k = 1, \dots, \omega$. A média desta distribuição F é μ , a média real de ninhos por tartaruga na população,

$$\mu = \sum_{k=1}^{\omega} k \alpha_k. \quad (5A.3)$$

Seja Y_j , $j = 1, \dots, T$, o número de ninhos da tartaruga j observados, $Y_j \in \{0, 1, 2, \dots, \omega\}$. Y_j é uma variável aleatória com distribuição binomial com parâmetros X_j e p :

$$Y_j \sim \text{Bin}(X_j, p) . \quad (5A.4)$$

Note-se que Y_j pode ser igual a zero no caso de uma tartaruga (que desovou) não ter nenhuma de suas desovas observada.

Seja G a distribuição de probabilidade dos $\{Y_j\}_{j=1}^T$,

$$Y_j \sim G = \{\beta_k\}_{k=0}^{\omega}, \quad \beta_k \geq 0, \quad \sum_{k=0}^{\omega} \beta_k = 1, \quad (5A.5)$$

onde $\beta_k = \#\{Y_j \mid Y_j = k\} / T$, ou seja, β_k é a proporção das tartarugas que tiveram k ninhos observados, $k = 0, 1, \dots, \omega$. Apenas os $Y_j > 0$ são observáveis, e portanto a distribuição G não pode ser calculada ou estimada diretamente a partir dos dados de campo.

Seja H a distribuição de probabilidade dos Y_j condicionados a $Y_j > 0$,

$$Y_j|_{Y_j>0} \sim H = \{\gamma_k\}_{k=1}^{\omega}, \quad \gamma_k \geq 0, \quad \sum_{k=1}^{\omega} \gamma_k = 1. \quad (5A.6)$$

A distribuição H é a distribuição esperada do número observado de ninhos entre as tartarugas observadas.

Se $\beta_0 < 1$, o que ocorre se $p > 0$, temos $\gamma_k = \beta_k / (\beta_1 + \dots + \beta_{\omega})$, $k = 1, \dots, \omega$, ou seja,

$$\gamma_k = \beta_k / (1 - \beta_0), \quad k = 1, \dots, \omega . \quad (5A.7)$$

Para cada tartaruga j , $j = 1, \dots, T$, para $k = 0, 1, \dots, \omega$, temos

$$\begin{aligned} \Pr(Y_j = k) &= \Pr\left(\bigcup_{i=1}^{\omega} (Y_j = k, X_j = i)\right) = \sum_{i=1}^{\omega} \Pr(Y_j = k, X_j = i) \\ &= \sum_{i=1}^{\omega} [\Pr(Y_j = k \mid X_j = i) \Pr(X_j = i)] = \sum_{i=1}^{\omega} \text{Bin}(i, p)(k) \alpha_i, \end{aligned} \quad (5A.8)$$

onde

$$\text{Bin}(i, p)(k) = \begin{cases} \binom{i}{k} p^k (1-p)^{i-k} & \text{se } 0 \leq k \leq i \\ 0 & \text{se } k > i \end{cases},$$

$$i = 1, \dots, \omega, \quad k = 0, 1, \dots, \omega . \quad (5A.9)$$

Como p é constante, pode ser adotada a notação

$$b_{ki} = \text{Bin}(i, p)(k), \quad i = 1, \dots, \omega, \quad k = 0, 1, \dots, \omega. \quad (5A.10)$$

Então,

$$\Pr(Y_j = k) = \sum_{i=1}^{\omega} b_{ki} \alpha_i, \quad k = 0, 1, \dots, \omega. \quad (5A.11)$$

Mas $\Pr(Y_j = k) = \beta_k$. Então,

$$\beta_k = \sum_{i=1}^{\omega} b_{ki} \alpha_i, \quad k = 0, 1, \dots, \omega. \quad (5A.12)$$

Substituindo os β_k da equação (5A.12) em (5A.7), temos, para $p > 0$,

$$\gamma_k = \sum_{i=1}^{\omega} b_{ki} \alpha_i / \left(1 - \sum_{i=1}^{\omega} b_{0i} \alpha_i \right), \quad k = 1, \dots, \omega, \quad (5A.13)$$

que é a própria equação (5.7).

CAPÍTULO 6

Propriedades estatísticas do modelo

As propriedades estatísticas do estimador apresentado no Capítulo 5 foram investigadas por meio de simulações em computador, utilizando-se as populações-modelo de tartarugas construídas segundo o método do Capítulo 3. No presente capítulo será mostrado que, em uma região definida de valores dos parâmetros ω e T , o estimador é essencialmente não-viciado (tem vício relativamente pequeno) e consistente. As simulações permitem também o cálculo do esforço de amostragem necessário para se obter um nível determinado de precisão na estimação de μ .

6.1. Metodologia de simulação

Cada simulação para o estudo das propriedades do estimador teve como base uma população-modelo de tartarugas, contida em um arquivo de dados. Foram realizadas 160 simulações diferentes, correspondentes a 160 populações-modelo, utilizando-se os arquivos de dados descritos no Capítulo 3, construídos segundo metodologia apresentada naquele capítulo. As simulações ocuparam no total cerca de 8 horas de tempo de processador no sistema IBM SP2. Para cada uma das populações-modelo (o que significa dizer: conhecida a distribuição real F do número de ninhos por tartaruga e a média desta distribuição, μ), a simulação

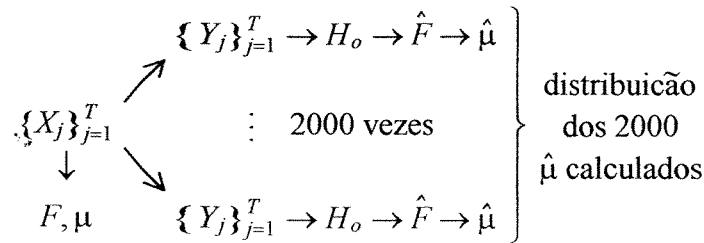


Figura 6.1. Simulação para cálculo da distribuição amostral de $\hat{\mu}$ para um valor de p fixo.

tem por objetivo a construção da distribuição amostral da estimativa $\hat{\mu}$ para cada valor fixo de p (a probabilidade de observação de cada ninho na praia), sendo que p pode assumir um número finito de valores (item 3 abaixo). Uma simulação, para um p fixo, reproduz repetidas vezes o processo de observação das tartarugas na praia em uma temporada e, para cada conjunto de dados pseudo-observados, calcula a estimativa $\hat{\mu}$, utilizando o modelo apresentado no Capítulo 5. As simulações foram feitas de acordo com o seguinte algoritmo (Figura 6.1):

1. Definir um valor para ω_e (ω para a estimação), o número máximo de ninhos por tartaruga que será admitido nas estimações da simulação. Este ω_e é o parâmetro designado por ω no modelo proposto no Capítulo 5. Em todas as simulações do presente capítulo, $\omega_e = \omega_d$, onde ω_d é definido no Capítulo 3.
2. Conhecido, do arquivo de dados, o número T de tartarugas e $\{X_j\}_{j=1}^T$, o número de ninhos de cada tartaruga da população, a distribuição F é calculada pelas equações (5A.1) e (5A.2), e a média μ desta distribuição F é calculada por (5A.3). (μ pode ser também calculada por $\mu = \sum_{j=1}^T X_j / T$.)

3. Para cada um dos seguintes valores de p : 0.98, 0.95, 0.90, 0.70, 0.50, 0.30, 0.20, 0.15, 0.10 e 0.05, repetir os itens 3.1-3.4 até que haja 2000 valores $\hat{\mu}$ calculados no item 3.4 ($\hat{\mu}$ é uma variável aleatória):

3.1. Para $j = 1, \dots, T$, gerar $Y_j \sim \text{Bin}(X_j, p)$ (equação (5A.4)), onde X_j é o número de ninhos da tartaruga j e Y_j é o número de seus ninhos observados.

3.2. Caso não exista nenhum $Y_j > 0$, retornar ao item 3.1. Se existir algum $Y_j > 0$, $j = 1, \dots, T$, para as tartarugas com $Y_j > 0$ calcular $H_o = (\gamma_1^o, \dots, \gamma_w^o)$, onde os γ_k^o são computados pela equação (5.5).

3.3. Calcular $\hat{p}$ por (5.2), onde $N_o = \sum_{Y_j > 0} Y_j$ e $N = \sum_{j=1}^T X_j$.

3.4. Calcular $\hat{F}$ por (5.12), utilizando-se $\hat{p}$ no cálculo da função f , e $\hat{\mu}$ por (5.13).

4. Da distribuição dos 2000 $\hat{\mu}$ calculados em 3.4 para cada valor fixo de p , computar medidas que resumem as propriedades desta distribuição para este valor de p : a média (M), o desvio-padrão (DP), o coeficiente de variação ($CV = DP/M$), o vício ($V = M - \mu$), o vício relativo ($VR = V/\mu$) e os percentis 2.5 e 97.5, que fornecem um intervalo de variação a 95% para a estimativa $\hat{\mu}$.

As Figuras 6.2-6.6 mostram os resultados das simulações para a distribuição RI7 e os cinco valores de T : 10, 20, 40, 100 e 250. Observe-se que, para $p = 1.0$, $\hat{\mu} = \mu$ sempre.

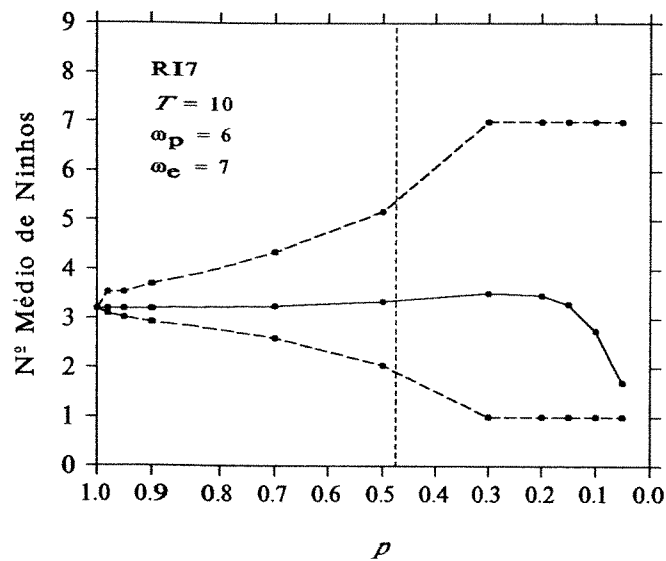


Figura 6.2. Número médio estimado de ninhos em função de p , resultados obtidos por simulação. Média (linha cheia) e percentis 2.5 e 97.5 (linhas tracejadas). Distribuição RI7, $T = 10$. Observar a escala reversa no eixo horizontal. A região útil do estimador está à esquerda da linha pontilhada vertical ($p \geq 0.475$).

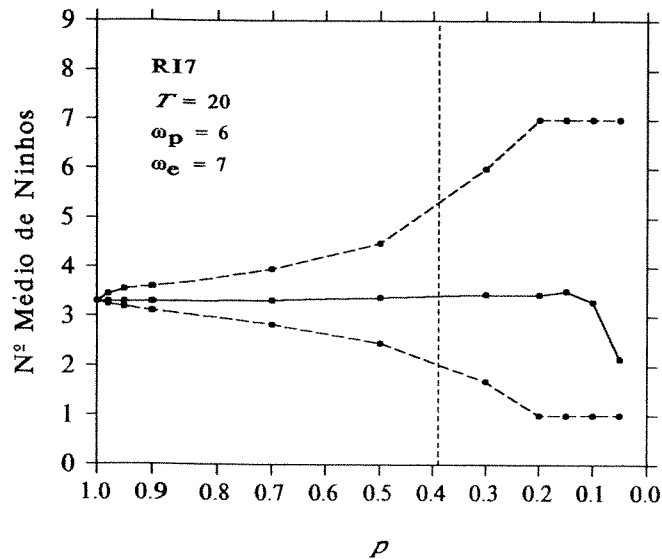


Figura 6.3. Número médio estimado de ninhos em função de p , resultados obtidos por simulação. Média (linha cheia) e percentis 2.5 e 97.5 (linhas tracejadas). Distribuição RI7, $T = 20$. Observar a escala reversa no eixo horizontal. A região útil do estimador está à esquerda da linha pontilhada vertical ($p \geq 0.389$).

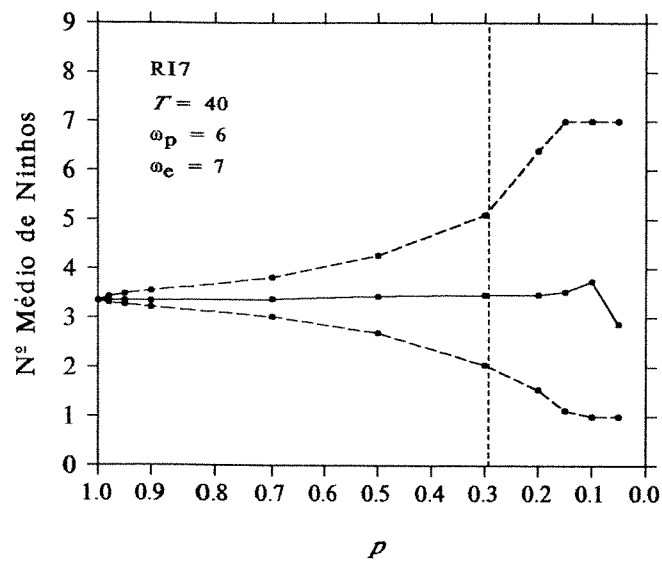


Figura 6.4. Número médio estimado de ninhos em função de p , resultados obtidos por simulação. Média (linha cheia) e percentis 2.5 e 97.5 (linhas tracejadas). Distribuição RI7, $T = 40$. Observar a escala reversa no eixo horizontal. A região útil do estimador está à esquerda da linha pontilhada vertical ($p \geq 0.292$).

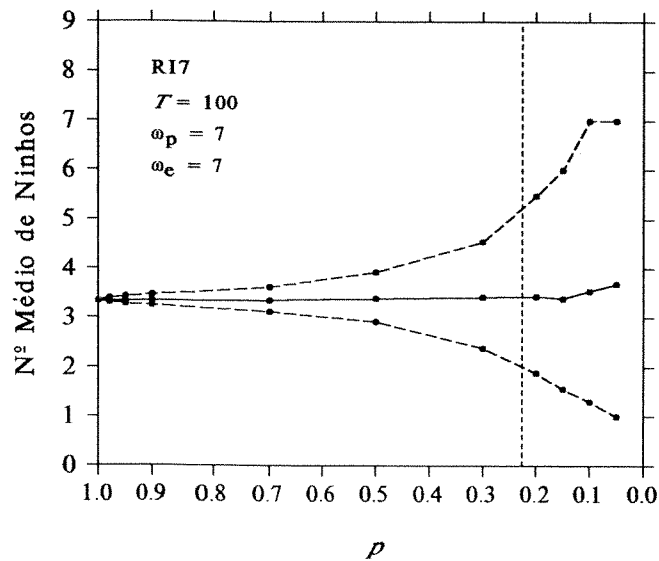


Figura 6.5. Número médio estimado de ninhos em função de p , resultados obtidos por simulação. Média (linha cheia) e percentis 2.5 e 97.5 (linhas tracejadas). Distribuição RI7, $T = 100$. Observar a escala reversa no eixo horizontal. A região útil do estimador está à esquerda da linha pontilhada vertical ($p \geq 0.227$).

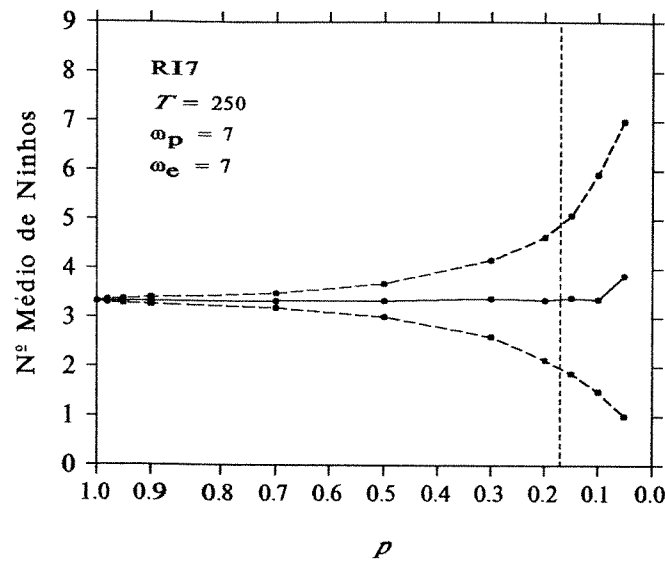


Figura 6.6. Número médio estimado de ninhos em função de p , resultados obtidos por simulação. Média (linha cheia) e percentis 2.5 e 97.5 (linhas tracejadas). Distribuição RI7, $T = 250$. Observar a escala reversa no eixo horizontal. A região útil do estimador está à esquerda da linha pontilhada vertical ($p \geq 0.170$).

6.2. Metodologia de análise dos resultados das simulações - a região útil do estimador

Como pode ser visto nas Figuras 6.2-6.6 (o padrão destas figuras é basicamente o mesmo para todas as outras distribuições estudadas), para uma distribuição-modelo e para um valor de T determinados, o estimador só fornece informação útil sobre μ em uma certa faixa de valores de p . Por informação útil entende-se informação adicional ao que já é conhecido sobre μ antes mesmo de ser tentada qualquer estimaco de seu valor real. Quando (item 1 do algoritmo de simulaco) é fixado o valor de ω_e , sabe-se, como decorrência, que μ está localizado entre 1, que é o número mínimo possível de ninhos por tartaruga, e ω_e , que é o número máximo de ninhos admitido como possível para uma tartaruga.

Ou seja, ao se estimar μ sabe-se *a priori* que $1 \leq \mu \leq \omega_e$. Se p for suficientemente pequeno, o processo de estimação pode não ser capaz de reduzir esta faixa de incerteza em relação ao valor de μ . Por exemplo, na Figura 6.2 (distribuição RI7, $T = 10$), vê-se que se $p < 0.30$ aproximadamente, o estimador não consegue dar qualquer informação sobre μ além do que já é conhecido *a priori*. Além disso, o processo de estimação na realidade não funciona para valores extremamente pequenos de p , pois quando $p \rightarrow 0$ pode ser mostrado que a média e os percentis da distribuição dos $\hat{\mu}$ tendem a 1, qualquer que seja o valor de μ (em relação a esta questão, veja a discussão no final do Capítulo 8). (Em termos práticos, um p menor dos que, digamos, cerca de 0.10 ou 0.05 provavelmente não existirá, pois se as desovas ocorrerem por exemplo durante cerca de 9 horas por noite, supondo uma distribuição temporal uniforme das desovas a cada noite, $p = 0.10$ e $p = 0.05$ corresponderão respectivamente a apenas cerca de 1 hora e a cerca de 30 minutos de trabalho de campo por noite).

Será denominada de *região útil* do estimador a faixa de valores de p que consegue dar informação adicional com interesse prático sobre o valor de μ além daquela conhecida *a priori*. Para a determinação desta região útil, o que envolve a definição de "interesse prático", será utilizado o critério apresentado em Krebs (1989, Capítulo 2), referente ao nível de precisão aceito para a estimativa: definindo-se o erro relativo E da estimativa por $E = (\hat{\mu} - \mu)/\mu$, deseja-se, com uma probabilidade (nível de confiança) de 0.95, que o erro relativo, em valor absoluto, seja igual ou menor que determinado valor. Serão determinados três níveis de interesse para o erro relativo, correspondentes a três situações de investigação sobre o valor de μ , cada nível de interesse fixando um valor $E_{\max}$ para o erro relativo máximo aceitável:

(1) Avaliação preliminar do valor de μ , quando se deseja apenas uma estimação grosseira do tamanho da população: $|E| \leq 0.50$, ou seja, $E_{\max} = 0.50$.

(2) Avaliação de μ para atividades de manejo, quando se deseja um nível moderado de erro: $|E| \leq 0.25$, ou seja, $E_{\max} = 0.25$.

(3) Avaliação de μ para trabalhos de pesquisa, quando o objetivo é obter uma estimativa relativamente próxima da realidade: $|E| \leq 0.10$, ou seja, $E_{\max} = 0.10$.

Embora nem sempre a distribuição dos 2000 valores de $\hat{\mu}$ calculados para cada p seja aproximadamente normal (na verdade, esta distribuição pode ser bastante assimétrica), se for aceita uma aproximação normal para esta distribuição o erro relativo máximo $E_{\max}$, a um nível de confiança de 0.95, pode ser aproximado por $E_{\max} = 2 DP / \mu$, onde DP é o desvio-padrão da distribuição para o valor de p considerado. Ou seja, $E_{\max} = 2 CV$, onde CV é o coeficiente de variação da distribuição para o p considerado, uma das medidas de resumo obtidas das simulações (item 4 do algoritmo da Seção 6.1). Assim, os três níveis de erro estabelecidos acima traduzem-se em três valores de interesse para o coeficiente de variação: $CV = 0.25$ (estimção grosseira), $CV = 0.125$ (estimção para manejo) e $CV = 0.05$ (estimção para pesquisa).

Para cada uma das 160 simulações, e para cada um dos três níveis para CV , foi então determinado o valor de p correspondente ao nível de CV considerado, ou seja, o valor de p , calculado por interpolação linear a partir dos pontos obtidos das simulações, para o qual a distribuição dos 2000 $\hat{\mu}$ tem um coeficiente de variação igual ao CV considerado. Estes três valores de p serão denotados por $p_{CV=0.25}$, $p_{CV=0.125}$ e $p_{CV=0.05}$. Para cada simulação, a região útil do estimador foi definida como a faixa de valores de p tal que $p \geq p_{CV=0.25}$, ou seja, a faixa de

valores de p suficientes para, no mínimo, uma estimação grosseira do valor de μ . Nas Figuras 6.2-6.6, $p_{CV=0.25}$, que delimita a região útil, está representado por uma linha pontilhada vertical.

Os cálculos realizados para a análise dos resultados das simulações foram realizados por programas escritos em Turbo Pascal 7.0 / DOS. Todas as análises sobre as propriedades estatísticas do modelo proposto no Capítulo 5 foram feitas apenas para valores de p dentro da região útil. Embora estas análises tenham sido feitas tendo por base o coeficiente de variação de $\hat{\mu}$, os resultados aplicam-se a $\hat{T}$, pois pode ser demonstrado que, no caso em que é válida a equação (5.1), quando N é conhecido exatamente, $CV_{\hat{T}} \cong CV_{\hat{\mu}}$, sendo que o erro relativo nesta aproximação é de cerca de 1% se $CV_{\hat{\mu}} = 0.05$, de cerca de 2% se $CV_{\hat{\mu}} = 0.125$ e de cerca de 6% se $CV_{\hat{\mu}} = 0.25$.

6.3. Resultados das simulações: propriedades do estimador

De acordo com os resultados obtidos nas simulações, o estimador definido no Capítulo 5 é, na região útil, essencialmente não-viciado (o vício, se existe, é relativamente pequeno) e consistente.

Estimador não-viciado. As Figuras 6.2-6.6 mostram que a média da distribuição dos 2000 $\hat{\mu}$ calculados para cada p , para a distribuição-modelo RI7, para os diferentes valores de T , é essencialmente igual a μ dentro da região útil, ou seja, o vício (V) do estimador é essencialmente igual a zero. As Figuras 6.7-6.9 resumem, para $CV = 0.05, 0.125$ e 0.25 , as características do estimador em relação ao vício para o conjunto das distribuições estudadas, para $5 \leq \omega_e \leq 8$. Estas figuras foram construídas de acordo com o seguinte método:

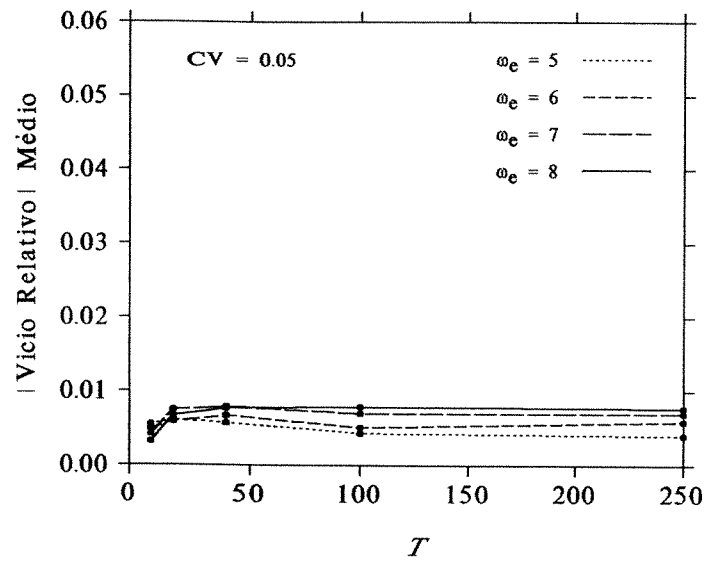


Figura 6.7. Vício relativo médio do estimador, em valor absoluto, em função de T , resultados obtidos por simulação para $5 \leq \omega_e \leq 8$. $CV = 0.05$.

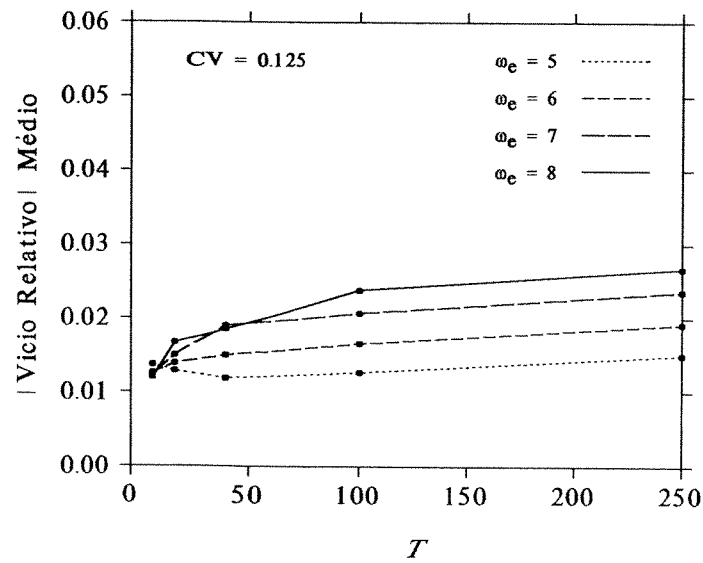


Figura 6.8. Vício relativo médio do estimador, em valor absoluto, em função de T , resultados obtidos por simulação para $5 \leq \omega_e \leq 8$. $CV = 0.125$.

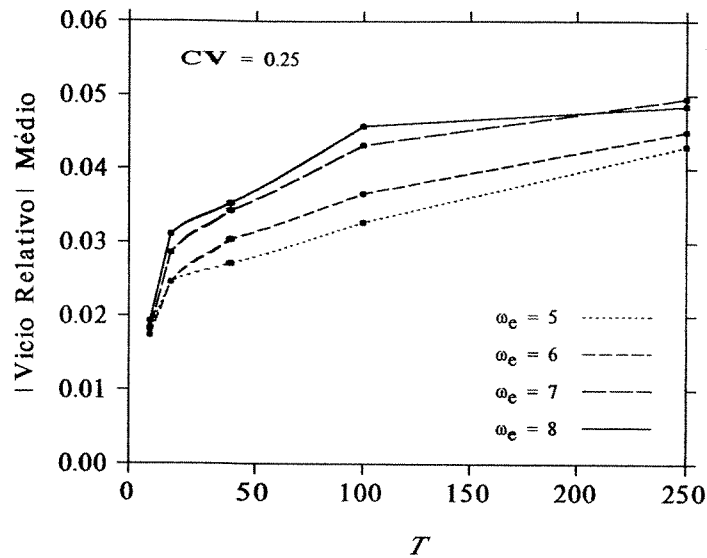


Figura 6.9. Vício relativo médio do estimador, em valor absoluto, em função de T , resultados obtidos por simulação para $5 \leq \omega_e \leq 8$. $CV = 0.25$.

1. Para cada uma das 160 distribuições estudadas, foram determinados $p_{CV=0.05}$, $p_{CV=0.125}$ e $p_{CV=0.25}$ (Seção 6.2).
2. Para cada distribuição, foram calculados $AVR_{CV=0.05}$, $AVR_{CV=0.125}$ e $AVR_{CV=0.25}$, os valores absolutos do vício relativo ($AVR = |V/\mu|$, é uma função de p) quando $p = p_{CV=0.05}$, $p = p_{CV=0.125}$ e $p = p_{CV=0.25}$, respectivamente. Foi feita uma interpolação linear entre os pontos obtidos diretamente das simulações, quando necessário.
3. Para cada valor de ω_e , $5 \leq \omega_e \leq 8$, para cada valor de T , $T = 10, 20, 40, 100$ e 250 , foram calculados $\overline{AVR}_{CV=0.05}(\omega_e, T)$ = média dos $AVR_{CV=0.05}$ computados no item (2) para as oito distribuições-modelo com parâmetros $\omega_d = \omega_e$ e T e, definidos de forma análoga, $\overline{AVR}_{CV=0.125}(\omega_e, T)$ e $\overline{AVR}_{CV=0.25}(\omega_e, T)$. Estes $\overline{AVR}_{CV}(\omega_e, T)$ médios válidos para cada conjunto de valores de CV , ω_e e T são plotados nas Figuras 6.7-6.9.

Observa-se nas Figuras 6.7-6.9 que o vício relativo, em valor absoluto, é em média menor que 0.01 para $CV = 0.05$ (estimação para pesquisa), menor que 0.03 para $CV = 0.125$ (estimação para manejo) e menor que 0.05 para $CV = 0.25$ (estimação grosseira). (O coeficiente de variação dos vícios relativos em valor absoluto referente às médias plotadas nas Figuras 6.7 - 6.9 é tipicamente da ordem de 0.50). Assim, mesmo no caso de uma estimação grosseira de μ o vício do estimador é, em valor absoluto, na região útil, em média no máximo 5% do valor de μ , para $5 \leq \omega_e \leq 8$ e $10 \leq T \leq 250$.

Estimador p -consistente. Para uma população finita, um estimador é consistente se, qualquer que seja o valor do parâmetro sendo estimado, a estimativa tende ao parâmetro sendo estimado quando o tamanho da amostra tende à população total. No caso do estimador proposto no Capítulo 5, para T fixo, o tamanho da amostra depende do parâmetro p . O estimador será definido como p -consistente se, qualquer que seja o valor de μ , para T fixo, $\hat{\mu} \rightarrow \mu$ quando $p \rightarrow 1$. As Figuras 6.10 e 6.11 resumem as propriedades do estimador apresentado no Capítulo 5 em relação à p -consistência, dentro da faixa de valores estudados para T , para $\omega_e = 5$. As figuras correspondentes para $\omega_e = 6, 7$ e 8 têm o mesmo padrão das figuras para $\omega_e = 5$ e não serão apresentadas. As Figuras 6.10 e 6.11 foram elaboradas de acordo com o seguinte método:

1. Para cada uma das 160 populações-modelo estudadas, foi determinado o valor $p_{CV=0.25}$ (Seção 6.2).
2. Para cada valor de ω_e , $5 \leq \omega_e \leq 8$, para cada valor de T , $T = 10, 20, 40, 100$ e 250 , foi calculado $p_{CV=0.25}^{\max}(\omega_e, T) = \max_{\text{distribuições com } \omega_e \text{ e } T} \{p_{CV=0.25}\}$.

Para uma população-modelo com parâmetros ω_e e T , quando

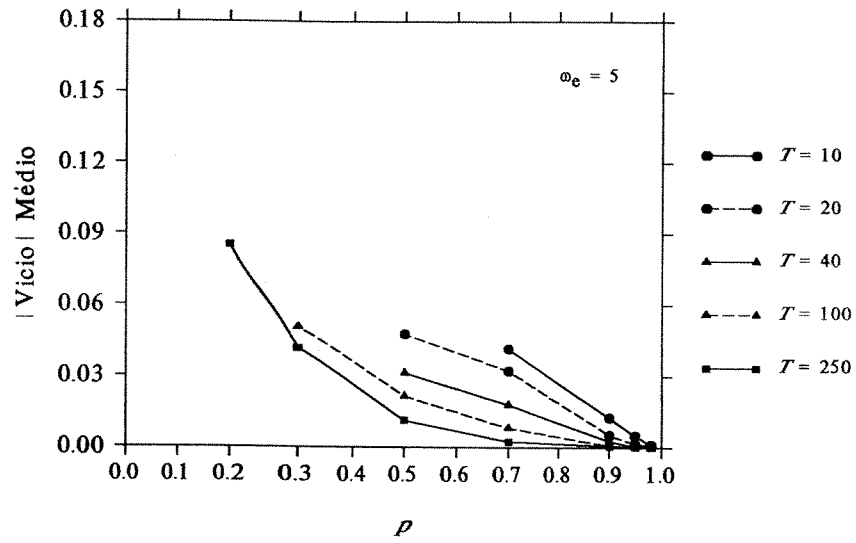


Figura 6.10. Vício médio do estimador, em valor absoluto, em função de p , para cada valor de T , resultados obtidos por simulação. $\omega_e = 5$.

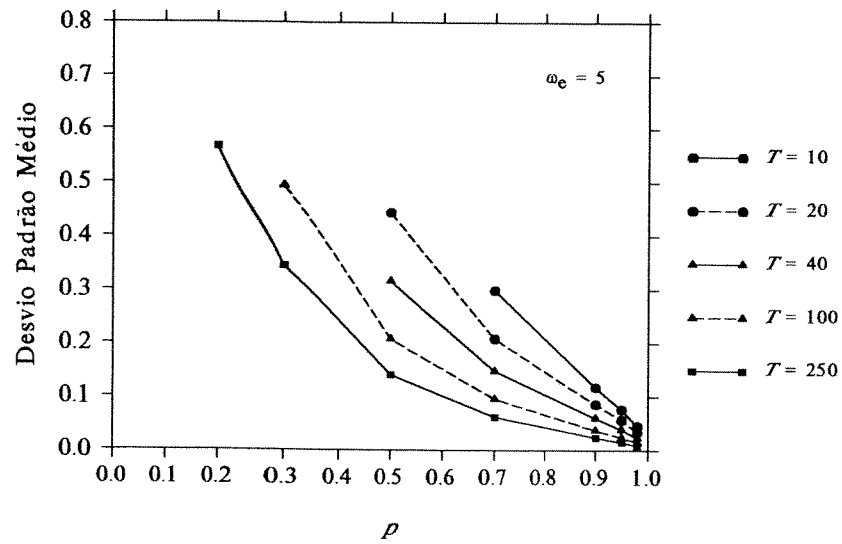


Figura 6.11. Desvio padrão médio da distribuição amostral de $\hat{\mu}$ gerada pelo estimador, em função de p , para cada valor de T , resultados obtidos por simulação. $\omega_e = 5$.

$p \geq p_{CV=0.25}^{\max}(\omega_e, T)$, o estimador, dentro do conjunto das 160 populações estudadas, estará certamente na região útil. Para todos os ω_e e T estudados, foi encontrado que $p_{CV=0.25}^{\max}(\omega_e, T) > 0.10$.

3. Para cada população-modelo, para cada valor de p , $p = 0.98, 0.95, 0.90, 0.70, 0.50, 0.30, 0.20, 0.15$ e 0.10 , quando $p \geq p_{CV=0.25}^{\max}(\omega_e, T)$, foram calculados (obtidos pelas simulações) o valor de DP (o desvio-padrão da distribuição amostral dos $\hat{\mu}$) e o valor de V (o vício desta distribuição amostral) para este p .
4. Para cada valor de ω_e , $5 \leq \omega_e \leq 8$, para cada valor de T , $T = 10, 20, 40, 100$ e 250 , para cada valor de p , foram computados (a) $\overline{DP}(\omega_e, T, p)$ = média dos DP referentes a p para as populações-modelo com parâmetros ω_e e T para as quais $p > p_{CV=0.25}^{\max}(\omega_e, T)$ e (b) $\overline{|V|}(\omega_e, T, p)$ = média dos valores absolutos dos V referentes a p também para as populações-modelo com parâmetros ω_e e T para as quais $p > p_{CV=0.25}^{\max}(\omega_e, T)$. A Figura 6.10 mostra $\overline{|V|}(\omega_e, T, p)$ para $\omega_e = 5$, e a Figura 6.11 mostra $\overline{DP}(\omega_e, T, p)$ também para $\omega_e = 5$. As figuras correspondentes para $\omega_e = 6, 7$ e 8 , como já foi dito, têm o mesmo padrão das Figuras 6.10 e 6.11.

As condições $\overline{DP} \rightarrow 0$ e $\overline{|V|} \rightarrow 0$ quando $p \rightarrow 1$, para T fixo, observadas nas Figuras 6.10 e 6.11, são suficientes para garantir a p -consistência do estimador (Casella e Berger, 1990, teorema 7.3.6 com $p \rightarrow 1$ no lugar de $n \rightarrow \infty$.)

6.4. Cálculo do esforço de amostragem

Os resultados das simulações realizadas com as 160 populações-modelo permitem que se obtenha uma estimativa do esforço de amostragem necessário

para se obter um nível especificado de precisão na estimação de μ . O esforço de amostragem será medido por p , a probabilidade de encontro de um ninho na praia. $p = 1.0$ corresponde a uma cobertura completa da praia pela equipe de campo. Um valor de p menor do que 1.0, que representa uma cobertura incompleta da praia, pode ser traduzido no esforço correspondente de trabalho de campo da seguinte forma: supondo que as tartarugas possam desovar a cada noite durante um período de h horas, para se obter um determinado valor de p no trabalho de campo será necessário que as atividades de campo tenham uma duração de h_o horas (h_o representa horas de observação), de tal forma que $h_o / h = p$. Como exemplo, se é admitido que as tartarugas possam desovar durante um período de 9 horas a cada noite, e se deseja um trabalho de campo produzindo $p = 0.75$, deveremos ter h_o tal que $h_o / 9 = 0.75$, ou seja, o trabalho de campo deverá ocupar 6.75 horas (6 horas e 45 minutos) a cada noite. Esta correspondência $p \leftrightarrow h_o$ pressupõe uma distribuição temporal uniforme das desovas a cada noite.

A Figura 6.12 mostra o valor aproximado de p necessário para a estimativa de μ com três diferentes possíveis níveis de precisão, especificados pelo valor de CV : $CV = 0.05$, $CV = 0.125$ e $CV = 0.25$. Esta figura foi construída de acordo com o seguinte método:

1. Para cada uma das 160 populações estudadas, foram determinados $p_{CV=0.05}$, $p_{CV=0.125}$ e $p_{CV=0.25}$ (Seção 6.2).
2. Para cada valor de ω_e , $5 \leq \omega_e \leq 8$, para cada valor de T , $T = 10, 20, 40, 100$ e 250 , foram calculados $\bar{p}_{CV=0.05}(\omega_e, T)$ = média dos $p_{CV=0.05}$ para as oito populações-modelo com parâmetros $\omega_d = \omega_e$ e T , e, definidos de forma análoga, $\bar{p}_{CV=0.125}(\omega_e, T)$ e $\bar{p}_{CV=0.25}(\omega_e, T)$. Estes valores médios

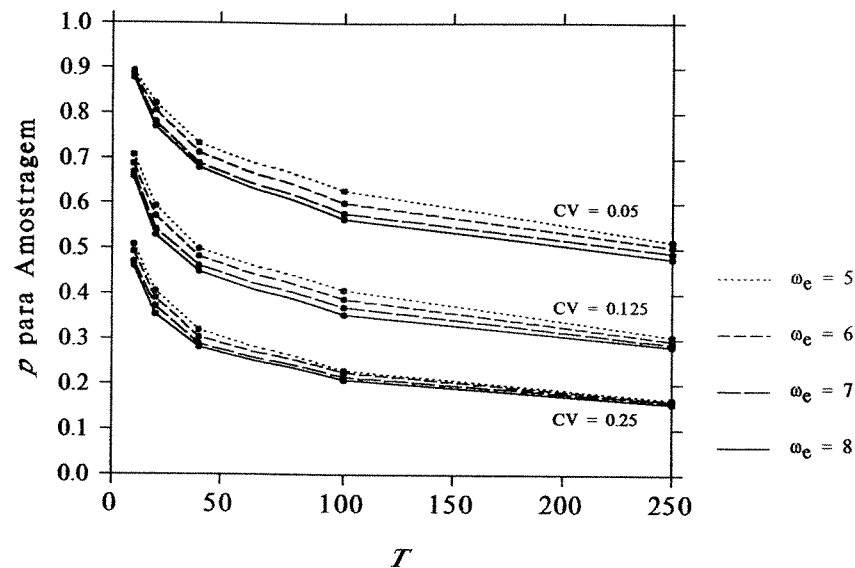


Figura 6.12. Esforço de amostragem, medido pela probabilidade p , necessário para a obtenção de um nível especificado de precisão na estimação de μ , medido por CV , em função de T , para $CV = 0.05, 0.125$ e 0.25 . Resultados obtidos por simulação para $10 \leq T \leq 250$ e $5 \leq \omega_e \leq 8$.

calculados, referentes aos três valores de CV , são plotados na Figura 6.12.

Para cada conjunto $\{CV, \omega_e, T\}$, o coeficiente de variação dos p_{CV} é tipicamente da ordem de 3%-10%, o que significa que os valores médios plotados na Figura 6.12 representam adequadamente o esforço de amostragem necessário para a obtenção de um nível de precisão especificado. A Figura 6.12 mostra que, para um nível de precisão fixo, à semelhança do que ocorre em outros esquemas de amostragem, o esforço de amostragem (no caso, a fração que os ninhos observados representam em relação ao total de ninhos) diminui à medida que T , o tamanho da população de tartarugas desovando na temporada, aumenta. Observa-se também na Figura 6.12 que o esforço de amostragem, medido por p , depende de forma bastante fraca do valor de ω_e , ao menos na

faixa $5 \leq \omega_e \leq 8$. O estudo da sensibilidade do estimador em relação a ω_e será realizado no Capítulo 7.

O uso da Figura 6.12 no planejamento do trabalho de campo necessário para a obtenção de um nível de precisão especificado na estimação de μ requer uma estimativa *a priori* para o valor de T . Um valor aproximado para T pode ser obtido com o uso da equação (4.1), utilizando-se como estimativa para μ , requerida por esta equação, algum valor obtido da literatura.

6.5. Discussão

O estimador proposto no Capítulo 5, conforme os resultados obtidos por meio das simulações, tem, na região útil, para $5 \leq \omega_e \leq 8$, boas propriedades estatísticas. A faixa de valores estudada para ω_e , $5 \leq \omega_e \leq 8$, como foi dito no Capítulo 3, é compatível com os valores máximos para o número de ninhos por tartaruga por temporada encontrados na literatura para as tartarugas *Caretta caretta* e *Chelonia mydas*. Para a aplicação do modelo a outras espécies, há necessidade de se verificar as propriedades do estimador para valores de ω_e menores do que 5 e maiores do que 8. Tartarugas marinhas do gênero *Lepidochelys* têm um número máximo de ninhos por tartaruga por temporada da ordem de 3 ou 4 (Reichart, 1993; Marquez, 1994), e para as *Dermochelys coriacea* existe registro de 11 ninhos por tartaruga em uma temporada (Boulon *et al.*, 1994).

CAPÍTULO 7

Análise de sensibilidade

A estimação de μ segundo o modelo proposto no Capítulo 5 requer que seja estipulado um valor para ω_e , o número máximo de ninhos por tartaruga admitido como possível em uma temporada. Neste capítulo será investigada, por meio de simulações em computador, a dependência da estimativa $\hat{\mu}$ em relação a ω_e . Será também discutida a sensibilidade do estimador em relação à perda de marcas, problema que aflige de forma crônica as pesquisas das tartarugas marinhas.

7.1. Sensibilidade em relação a ω_e : metodologia de simulação

O objetivo das simulações para a análise da sensibilidade do modelo em relação a ω_e é comparar, para diferentes distribuições observadas H_o , as estimativas $\hat{\mu}_{\omega_e}$ obtidas a partir de cada H_o utilizando-se diferentes valores de ω_e no estimador do Capítulo 5. Cada simulação teve como base uma população-modelo de tartarugas, contida em um arquivo de dados construído conforme descrito no Capítulo 3. Foram realizadas 48 simulações, agrupadas da seguinte forma:

1. Para $T = 10$, foram estudadas as distribuições AA6, AY5, BC5, BM6,

FL5, RI5, RI6, SR7, SR8 e ST6. Os arquivos de dados referentes a estas 10 distribuições, para $T = 10$, têm $\omega_p \leq 5$ (ω_p é definido na Seção 3.4), e as simulações foram feitas com $\omega_e = 5, 6, 7$ e 8 .

2. Para $T = 20$, foram estudadas as distribuições AA6, AY5, BC5, BM5, FL5, RI5, RI6, SR5, SR6 e ST5. Os arquivos de dados referentes a estas 10 distribuições, para $T = 20$, têm $\omega_p \leq 5$, e as simulações foram feitas com $\omega_e = 5, 6, 7$ e 8 .

3. Para $T = 40$, foram estudadas as distribuições AA6, AY5, BC5, BC6, BM5, FL5, RI5, RI6, SR5 e ST5. Os arquivos de dados referentes a estas 10 distribuições, para $T = 40$, têm $\omega_p \leq 5$, e as simulações foram feitas com $\omega_e = 5, 6, 7$ e 8 .

4. Para $T = 100$, foram estudadas as distribuições AA5, AY5, BC5, BM5, FL5, RI5, SR5 e ST5. Os arquivos de dados referentes a estas 8 distribuições, para $T = 100$, têm $\omega_p \leq 5$, e as simulações foram feitas com $\omega_e = 5, 6, 7$ e 8 .

5. Para $T = 250$, foram estudadas as distribuições AA6, AY5, BC5, BM6, FL5, RI5, RI6, SR5, SR6 e ST6. Os arquivos de dados referentes a estas 10 distribuições, para $T = 250$, têm $\omega_p \leq 6$, e as simulações foram feitas com $\omega_e = 6, 7$ e 8 .

As simulações ocuparam no total cerca de 1 hora de tempo de processador no sistema IBM SP2. Cada simulação foi feita de acordo com o seguinte algoritmo:

1. Conhecido, do arquivo de dados, o número T de tartarugas e $\{X_j\}_{j=1}^T$, onde X_j é o número de ninhos da tartaruga j , para cada um dos seguintes

valores de p : 1.00, 0.98, 0.95, 0.90, 0.70, 0.50, 0.30, 0.20 e 0.15, repetir os itens 1.1-1.4 até que haja 400 coeficientes de variação (CV 's) calculados no item 1.4 (CV é uma variável aleatória):

1.1. Para $j = 1, \dots, T$, gerar $Y_j \sim \text{Bin}(X_j, p)$ (equação (5A.4)), onde Y_j é o número de ninhos observados da tartaruga j .

1.2. Caso não exista nenhum $Y_j > 0$, $j = 1, \dots, T$, retornar ao item 1.1. Se existir algum $Y_j > 0$, para as tartarugas com $Y_j > 0$ calcular

$H_o = (\gamma_1^o, \dots, \gamma_\omega^o)$, onde γ_k^o é computado pela equação (5.5).

1.3. Para cada valor de ω_e , $\omega_e = 5, 6, 7$ e 8 (ou apenas $\omega_e = 6, 7$ e 8 , se $T = 250$), calcular $\hat{F}_{\omega_e}$ por (5.12) e $\hat{\mu}_{\omega_e}$ por (5.13), utilizando para todos os valores de ω_e o mesmo H_o computado no item 1.2.

1.4. Calcular o coeficiente de variação ($CV = \text{desvio-padrão} / \text{média}$) entre os $\hat{\mu}_{\omega_e}$ estimados no item 1.3.

2. Da distribuição dos 400 CV calculados para cada p , calcular a média e os percentis 2.5 e 97.5.

O percentil 2.5 foi em todas as simulações, para todos os valores de p , igual a zero. A média e o percentil 97.5, como funções de p , estão plotados nas Figuras 7.1 - 7.5 para cada conjunto de distribuições estudado. Nestas figuras, cada população-modelo está representada por duas linhas: uma linha cheia (a média dos CV para esta população) e uma linha tracejada (o percentil 97.5 dos CV). Assim, por exemplo, na Figura 7.1 há 10 linhas cheias, representando a média dos CV para as 10 distribuições-modelo estudadas para $T = 10$, e 10 linhas tracejadas, referentes aos percentis 97.5 dos CV para estas 10 distribuições-modelo.

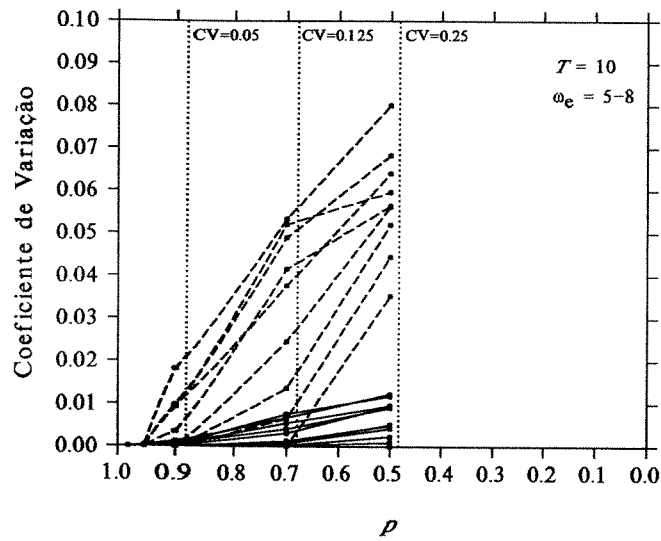


Figura 7.1. Média (linha cheia) e percentil 97.5 (linha tracejada) do coeficiente de variação dos $\hat{\mu}_{\omega_e}$ em função de p para $5 \leq \omega_e \leq 8$, resultados obtidos por simulação para $T = 10$. Cada população-modelo (veja o texto) tem os seus resultados representados por duas linhas, uma cheia e uma tracejada. As linhas pontilhadas verticais indicam $\bar{p}_{CV}(10)$ para $CV = 0.05, 0.125$ e 0.25 . Observar a escala reversa no eixo horizontal.

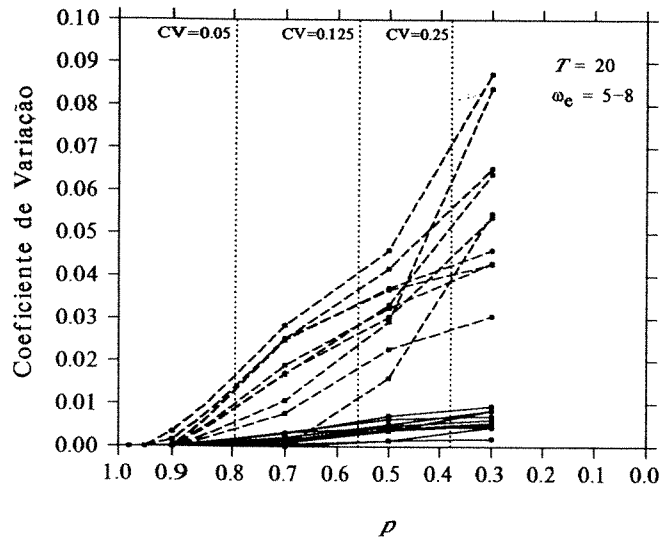


Figura 7.2. Média (linha cheia) e percentil 97.5 (linha tracejada) do coeficiente de variação dos $\hat{\mu}_{\omega_e}$ em função de p para $5 \leq \omega_e \leq 8$, resultados obtidos por simulação para $T = 20$. Cada população-modelo (veja o texto) tem os seus resultados representados por duas linhas, uma cheia e uma tracejada. As linhas pontilhadas verticais indicam $\bar{p}_{CV}(20)$ para $CV = 0.05, 0.125$ e 0.25 . Observar a escala reversa no eixo horizontal.

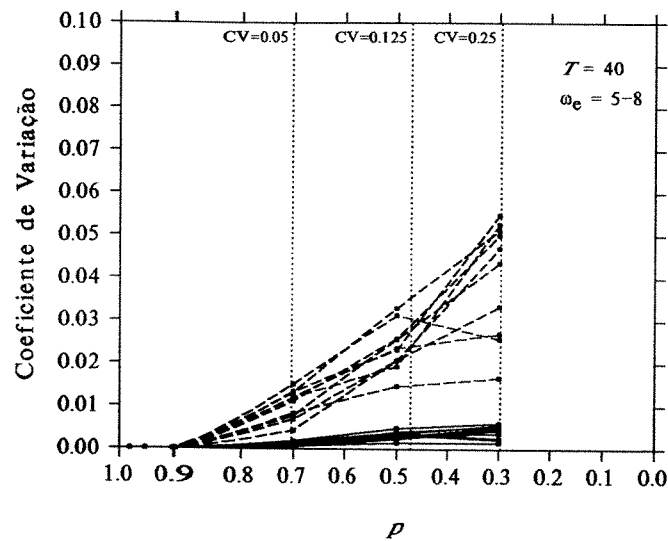


Figura 7.3. Média (linha cheia) e percentil 97.5 (linha tracejada) do coeficiente de variação dos $\hat{\mu}_{\omega_e}$ em função de ρ para $5 \leq \omega_e \leq 8$, resultados obtidos por simulação para $T = 40$. Cada população-modelo (veja o texto) tem os seus resultados representados por duas linhas, uma cheia e uma tracejada. As linhas pontilhadas verticais indicam $\bar{p}_{CV}(40)$ para $CV = 0.05, 0.125$ e 0.25 . Observar a escala reversa no eixo horizontal.

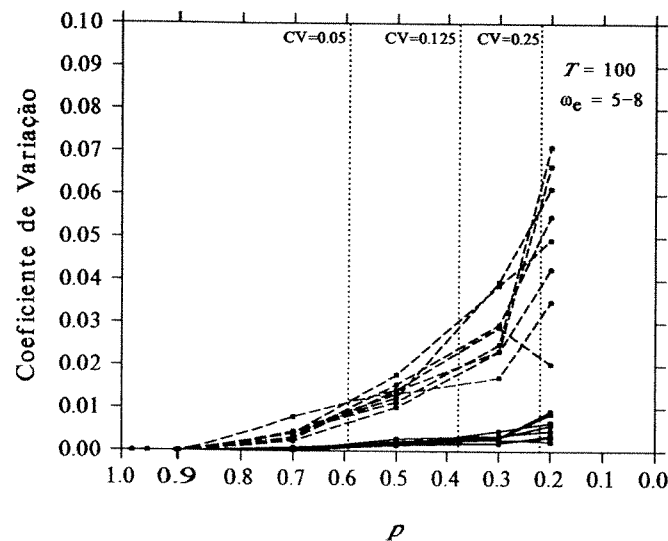


Figura 7.4. Média (linha cheia) e percentil 97.5 (linha tracejada) do coeficiente de variação dos $\hat{\mu}_{\omega_e}$ em função de ρ para $5 \leq \omega_e \leq 8$, resultados obtidos por simulação para $T = 100$. Cada população-modelo (veja o texto) tem os seus resultados representados por duas linhas, uma cheia e uma tracejada. As linhas pontilhadas verticais indicam $\bar{p}_{CV}(100)$ para $CV = 0.05, 0.125$ e 0.25 . Observar a escala reversa no eixo horizontal.

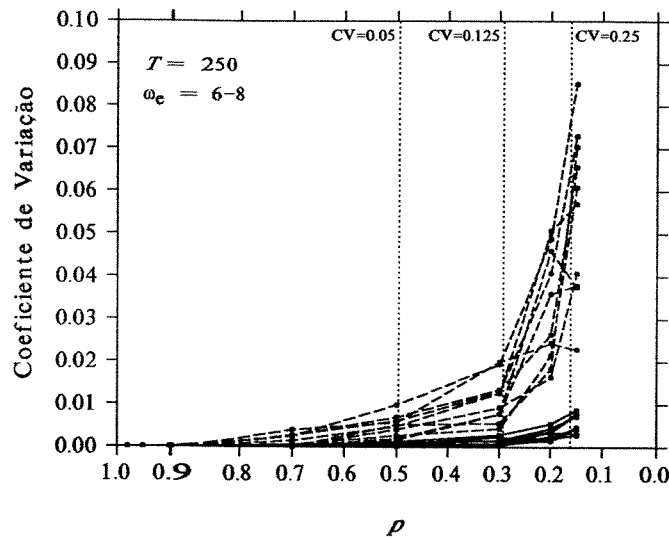


Figura 7.5. Média (linha cheia) e percentil 97.5 (linha tracejada) do coeficiente de variação dos $\hat{\mu}_{\omega_e}$ em função de p para $6 \leq \omega_e \leq 8$, resultados obtidos por simulação para $T = 250$. Cada população-modelo (veja o texto) tem os seus resultados representados por duas linhas, uma cheia e uma tracejada. As linhas pontilhadas verticais indicam $\bar{p}_{CV}(250)$ para $CV = 0.05, 0.125$ e 0.25 . Observar a escala reversa no eixo horizontal.

Nas Figuras 7.1-7.5, as linhas pontilhadas verticais indicam os valores

$\bar{p}_{CV=0.05}(T)$, $\bar{p}_{CV=0.125}(T)$ e $\bar{p}_{CV=0.25}(T)$ para as distribuições com um mesmo valor de T . Para um valor fixo de T , $\bar{p}_{CV=0.05}(T)$ é a média dos $p_{CV=0.05}$ (definidos no Capítulo 6) para as populações finitas com este valor de T , qualquer que seja a população-modelo e qualquer que seja ω_d , $5 \leq \omega_d \leq 8$. Os valores $\bar{p}_{CV=0.125}(T)$ e $\bar{p}_{CV=0.25}(T)$ são definidos de forma análoga. Estas quantidades, $\bar{p}_{CV=0.05}(T)$, $\bar{p}_{CV=0.125}(T)$ e $\bar{p}_{CV=0.25}(T)$, são tomadas como aproximações aos valores $p_{CV=0.05}$, $p_{CV=0.125}$ e $p_{CV=0.25}$ para uma população finita com T tartarugas, qualquer que seja ω_d , $5 \leq \omega_d \leq 8$ e qualquer que seja a distribuição-modelo da população. A Figura 6.12 mostra que, para um mesmo valor de T , os p_{CV} dependem fracamente de ω_e , para $5 \leq \omega_e \leq 8$, o que justifica usar a média geral $\bar{p}_{CV}(T)$ como valor representativo de p_{CV} para um valor fixo de T .

7.2. Sensibilidade em relação a ω_e : resultados

As Figuras 7.1 - 7.5 mostram que, para $10 \leq T \leq 250$, para $CV = 0.05$ (estimação para pesquisa), a estimativa praticamente não depende da escolha de ω_e , para $5 \leq \omega_e \leq 8$. Para $CV = 0.125$ (estimação para manejo), as figuras mostram que os CV são em média menores que 0.01 e em geral (conforme o percentil 97.5) menores que 0.06. Para $CV = 0.25$ (estimação grosseira), os CV são em média menores que 0.01, e são em geral menores que 0.09. Observa-se assim que, principalmente nas estimações para pesquisa e para manejo, para $5 \leq \omega_e \leq 8$, a dependência da estimativa $\hat{\mu}$ em relação a ω_e é bastante fraca, ou, em outros termos, o estimador é bastante robusto em relação à escolha de ω_e .

7.3. O problema da perda de marcas

A perda de marcas é um problema recorrente nas pesquisas das tartarugas marinhas (Carr, 1980; Hughes, 1982; Frazer, 1983; Mrosovsky, 1983; Limpus, 1985, 1992b). Atualmente utiliza-se com frequência um sistema de marcação dupla, em que cada tartaruga porta duas marcas diferentes, o que tende a reduzir a dimensão do problema (Limpus, 1992b). Não obstante, o problema continua existindo, e uma questão que se coloca a um método de estimação de população que utilize dados obtidos por meio de marcação é avaliar o peso que a perda de marcas tem no processo de estimação considerado e procurar formas de correção do erro introduzido pela perda de marcas, se for o caso.

A perda de marcas leva, em princípio, a uma superestimação do número de tartarugas ocorrendo em uma temporada, pois uma tartaruga que tenha perdido suas marcas, se for novamente encontrada na praia na mesma temporada, será considerada como sendo uma nova tartaruga, recebendo novas marcas que não

terão como ser relacionadas às marcas perdidas. Em relação à estimação do número de tartarugas desovando em uma temporada através do uso do estimador clássico (Equação 4.1), esta superestimação ocorre devido a uma subestimação de μ , pois uma mesma tartaruga aparecerá nos dados como duas tartarugas diferentes, fictícias, cada uma destas com uma produção de ninhos na temporada inferior ao número de ninhos da tartaruga real.

Limpus (1992b), com base em pesquisas realizadas na Austrália durante um período de tempo superior a 10 anos, apresenta dados, válidos para estudos em praias de desova, mostrando a probabilidade da perda de uma marca em função do tempo decorrido desde sua aplicação. Para as marcas-padrão confeccionadas em monel e aplicadas em tartarugas *Caretta caretta* na região proximal das nadadeiras anteriores (posições L3 e R3 em Limpus (1992b)), a probabilidade de perda de uma marca varia aproximadamente linearmente com o tempo, chegando a 1.0 em um período de cerca de 10 anos após a aplicação (a variação linear no tempo é um modelo talvez grosseiro para os dados de Limpus (1992b) relativos à marcação de tartarugas da espécie *Chelonia mydas*). Ou seja, a probabilidade de perda de uma marca como função do tempo, $\pi(t)$, com o tempo medido em dias, pode ser representada por

$$\pi(t) = \frac{1}{10 \times 365} t, \quad t \leq 10 \times 365 \text{ dias.} \quad (7.1)$$

Marcas confeccionadas em titânio, utilizadas atualmente nas pesquisas na Austrália, aplicadas na mesma posição das nadadeiras anteriores, têm probabilidade de perda, para um mesmo t , inferior à das marcas em monel (Limpus, 1992b). Ou seja, segundo os dados de Limpus, a probabilidade de perda das marcas em monel (equação (7.1)) pode ser tomada como um limite superior

para a probabilidade da perda de uma marca aplicada (com técnicas adequadas) na região proximal das nadadeiras anteriores, em estudos realizados em praias de desova.

Em um processo de marcação dupla, com uma marca colocada em cada uma das nadadeiras anteriores da tartaruga (Limpus, 1992b), se considerarmos as marcas como independentes, a probabilidade de perda de ambas as marcas em função do tempo, $\pi_2(t)$, com o tempo medido em dias, será dada por

$$\begin{aligned}\pi_2(t) &= \pi(t) \times \pi(t) = \left(\frac{1}{10 \times 365}\right)^2 t^2 \\ &= (7.51 \times 10^{-8}) t^2, \quad t \leq 10 \times 365 \text{ dias.}\end{aligned}\tag{7.2}$$

Para uma tartaruga (a exemplo das *Caretta caretta*) que desova um máximo de cerca de 7 vezes em uma temporada, com um intervalo entre desovas de cerca de 14 dias, a probabilidade de perda das duas marcas na temporada será menor que $7.51 \times 10^{-8} \times (6 \times 14)^2 = 0.00053$, o que é uma probabilidade extremamente pequena. Seria necessário, em média, que se marcassem pelo menos $1.0 / 0.00053 \cong 1890$ tartarugas para que houvesse um caso de perda de ambas as marcas na temporada. Mesmo que a probabilidade de perda de uma marca, para cada t , fosse o dobro daquela estabelecida pela equação (7.1) (ou seja, a probabilidade de perda chegaria a 1.0 em 5 anos), seria necessário, em média, que se marcassem no mínimo cerca de 470 tartarugas para que, em um sistema de marcação dupla, houvesse um caso de perda de ambas as marcas, supondo-se independência na perda das duas marcas. Deve ser observado que esta hipótese de independência na perda das duas marcas pode não ser obedecida, como por exemplo na situação em que as duas marcas são retiradas em tartarugas presas em redes de pesca (Frazer, 1983).

Se p , a medida do esforço de amostragem na praia, tiver um valor razoavelmente alto, dificilmente uma tartaruga que desova mais de uma vez na temporada deixará de ser encontrada entre a primeira e a última desova, o que tornaria necessário que se marcasse um número de tartarugas superior (possivelmente bastante superior) àqueles valores mínimos calculados para que houvesse um caso de perda das duas marcas, uma vez que cada encontro de uma tartaruga possibilita que se substitua marcas em mal estado e se complete o seu par de marcas, caso uma das marcas tenha se perdido.

Os cálculos realizados mostram que, com um sistema de marcação dupla, com as técnicas de marcação utilizadas atualmente nas pesquisas das tartarugas marinhas, a perda de marcas terá provavelmente um efeito extremamente pequeno em um método de estimação do número de tartarugas válido para uma temporada de desova, principalmente se p tiver um valor razoavelmente alto.

CAPÍTULO 8

Construção de um intervalo de confiança

Um intervalo de confiança para μ , com a estimação feita segundo o modelo do Capítulo 5, será construído com a aplicação do método do *bootstrap*, um dos chamados métodos de reamostragem (Efron, 1982; Efron e Tibshirani, 1993). Será utilizado o método do *bootstrap* BCa (Efron e Tibshirani, 1993, Capítulo 14), com algumas adaptações, seguido de um processo de calibração. O método aqui proposto para a construção do intervalo de confiança foi avaliado por meio de simulações em computador, aplicando-se todo o processo observação-estimação-*bootstrap* a populações-modelo construídas segundo o método do Capítulo 3.

8.1. O intervalo de confiança

Dado $\hat{p} > 0$, onde $\hat{p}$ é a probabilidade estimada de observação de um ninho (equação 5.2), o estimador definido por (5.12) produz, a partir de uma distribuição observada H_o , uma estimativa $\hat{F}$ da distribuição F do número real de ninhos por tartaruga. Com base na distribuição estimada $\hat{F}$ são calculados $\hat{\mu}$ (por (5.13)) e $\hat{T}$ (por (5.1)).

Pelo método do *bootstrap* (Figura 8.1), será feita uma reconstrução (aproximada) do processo amostral real que gera a seqüência $F \rightarrow H_o \rightarrow \hat{F} \rightarrow \hat{\mu}$.

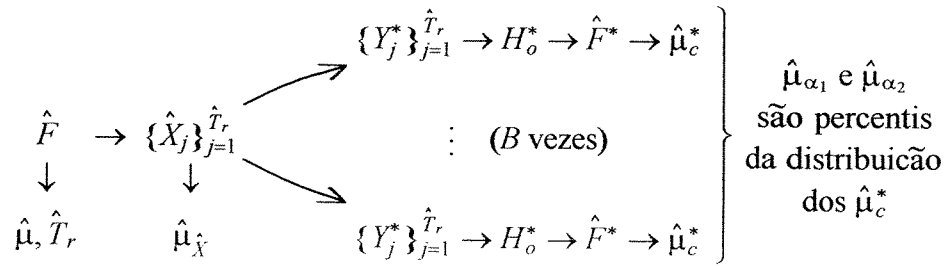


Figura 8.1. Construção de um intervalo de confiança pelo método do *bootstrap*. No *bootstrap* BCa (Efron e Tibshirani, 1993), α_1 e α_2 são correções respectivamente a α e $1 - \alpha$, de modo que o intervalo $(\hat{\mu}_{\alpha_1}, \hat{\mu}_{\alpha_2})$ tenha cobertura aproximada $1 - 2\alpha$.

A distribuição estimada $\hat{F}$ será tomada como sendo a distribuição real F , e se fará repetidas vezes, sempre com o mesmo $\hat{p}$, o mesmo usado em (5.12), uma simulação do processo observação-estimação a partir de uma mesma população finita com $\hat{T}$ tartarugas, população esta construída de forma que o número de ninhos por tartaruga siga aproximadamente (segundo método descrito no Capítulo 3) a distribuição $\hat{F}$. Será gerada assim uma distribuição amostral tomada como aproximação para a distribuição amostral de $\hat{\mu}$, e o intervalo de confiança para μ será construído com o uso de percentis desta distribuição amostral aproximada. Como $\hat{T}$ não é em geral um número inteiro, será arredondado para o número inteiro mais próximo, $\hat{T}_r$, antes de se iniciar o *bootstrap*.

Dado α , $0 < \alpha < 0.5$, o objetivo é construir um intervalo de confiança $(\hat{\mu}_{\alpha_1}, \hat{\mu}_{\alpha_2})$ tal que (1) $\Pr\{\hat{\mu}_{\alpha_1} \leq \mu \leq \hat{\mu}_{\alpha_2}\} = 1 - 2\alpha$ e (2) $\Pr\{\mu < \hat{\mu}_{\alpha_1}\} = \alpha$ e $\Pr\{\mu > \hat{\mu}_{\alpha_2}\} = \alpha$. O item (2) diz que o erro de cobertura do intervalo deverá ser dividido igualmente entre os extremos do intervalo construído. Os valores de α

usualmente utilizados são $\alpha = 0.05$ ou $\alpha = 0.025$, gerando respectivamente intervalos de confiança a 90% e a 95%.

A construção do intervalo de confiança por um *bootstrap* BCa será feita pelo seguinte algoritmo (Figura 8.1), pressupondo-se que já se tenha calculado as estimativas $\hat{F}$, $\hat{\mu}$ e $\hat{T}$:

1. Dados iniciais:

$\hat{p}$: a probabilidade estimada de observação de cada ninho, a mesma utilizada na estimação de $\hat{F}$ por (5.12).

N : o número total de ninhos colocados na temporada pelo conjunto das tartarugas. É o mesmo N utilizado na estimação de $\hat{F}$, $\hat{\mu}$ e $\hat{T}$.

ω : o número máximo de desovas que pode ser feito por uma tartaruga na temporada. É o mesmo ω utilizado na estimação de $\hat{F}$, $\hat{\mu}$ e $\hat{T}$.

$\hat{\mu}$: o número médio estimado de ninhos por tartaruga, calculado por (5.13).

$\hat{T}_r$: o número inteiro estimado de tartarugas. Para se obter $\hat{T}_r$, arredonda-se o $\hat{T}$ calculado por (5.1) para o número inteiro mais próximo.

$\hat{F}$: a distribuição estimada para o número de ninhos por tartaruga, calculada por (5.12).

α : parâmetro definindo o intervalo de confiança desejado, $0 < \alpha < 0.5$.

B : o número de amostras *bootstrap* a serem computadas. Para o método do *bootstrap* por percentis, do qual o *bootstrap* BCa é derivado, B deve ser maior que 500 ou 1000, de forma a se obter uma variabilidade relativamente pequena nos percentis calculados (Efron e Tibshirani, 1993, Capítulo 19). Nos *bootstraps* realizados no presente trabalho, foi adotado $B = 1500$.

2. Construir uma população $\{\hat{X}_j\}_{j=1}^{\hat{T}_r}$ com $\hat{T}_r$ tartarugas, onde $\hat{X}_j$ é o número de ninhos colocados pela tartaruga j , $j = 1, \dots, \hat{T}_r$, $\hat{X}_j \in \{1, \dots, \omega\}$, com $\{\hat{X}_j\}_{j=1}^{\hat{T}_r}$ seguindo o mais próximo possível a distribuição $\hat{F}$, utilizando-se a mesma metodologia empregada na construção das populações-modelo, descrita no Capítulo 3.

3. Calcular $\hat{\mu}_{\hat{X}} = \sum_{j=1}^{\hat{T}_r} \hat{X}_j / \hat{T}_r$, a média dos $\hat{X}_j$, e calcular $erro = \hat{\mu}_{\hat{X}} - \hat{\mu}$.

4. Fazer $b \leftarrow 0$ (b é o contador das amostras *bootstrap*) e executar os itens 4.1 - 4.4 até que haja B valores $\hat{\mu}_c^*$ calculados no item 4.4 ($\hat{\mu}_c^*$ é uma variável aleatória). Variáveis marcadas com o sinal * fazem parte de uma amostra *bootstrap*:

4.1. Gerar $Y_j^* \sim \text{Bin}(\hat{X}_j, \hat{p})$, $j = 1, \dots, \hat{T}_r$, Y_j^* representando o número observado de ninhos para a tartaruga j .

4.2. Caso não exista nenhum $Y_j^* > 0$, $j = 1, \dots, \hat{T}_r$, retornar ao item 4.1. Se existir algum $Y_j^* > 0$, para as tartarugas com $Y_j^* > 0$ calcular $H_o^* = \{\gamma_k^{o*}\}_{k=1}^{\omega}$, a distribuição do número de ninhos por tartaruga entre as tartarugas observadas, onde γ_k^{o*} é computado por (5.5). Fazer $b \leftarrow b + 1$.

4.3. Calcular $\hat{F}^* = \{\hat{\alpha}_k^*\}_{k=1}^{\omega}$ aplicando (5.12) a H_o^* , com $\hat{p}$ sendo utilizado no cálculo de f em (5.12).

4.4. Calcular $\hat{\mu}^*$ aplicando (5.13) a $\hat{F}^* = \{\hat{\alpha}_k^*\}_{k=1}^{\omega}$, e calcular a estimativa corrigida $\hat{\mu}_c^*(b) = \hat{\mu}^* - erro$, onde $erro$ foi calculado no item 3.

5. Calcular $\hat{\mu}_{\alpha_1} = \text{percentil } \alpha_1$ e $\hat{\mu}_{\alpha_2} = \text{percentil } \alpha_2$ da distribuição dos B valores $\hat{\mu}_c^*(b)$ calculados no item 4.4, onde α_1 e α_2 são calculados pela fórmula 14.10 de Efron e Tibshirani (1993), de acordo com o método do

bootstrap BCa. Neste *bootstrap* BCa, a constante $\hat{a}$ do *bootstrap* será computada de forma não-paramétrica pela fórmula 14.15 de Efron e Tibshirani (1993), e a constante $\hat{z}_0$ será computada pela fórmula 14.14 destes autores.

6. O intervalo de confiança $1 - 2\alpha$ será $(\hat{\mu}_{\alpha_1}, \hat{\mu}_{\alpha_2})$.

Para a situação em que o modelo apresentado no Capítulo 5 é aplicável, quando N , o número total de desovas feitas na temporada, é conhecido exatamente, obter um intervalo de confiança a um nível $1 - 2\alpha$ para T é equivalente a obter um intervalo de confiança a este mesmo nível para μ , uma vez que pode ser demonstrado que, qualquer que seja β , $0 < \beta < 1$, se $\hat{\mu}_\beta$ é o percentil β da distribuição $\{\hat{\mu}_c^*(b) \mid b = 1, \dots, B\}$, então $N / \hat{\mu}_\beta$ é o percentil $1 - \beta$ da distribuição $\{\hat{T}^*(b) \mid \hat{T}^*(b) = N / \hat{\mu}_c^*(b), b = 1, \dots, B\}$. Por este motivo, o *bootstrap* é discutido apenas em relação à estimativa $\hat{\mu}$.

8.2. Discussão sobre o algoritmo

Duas aproximações estão sendo feitas no algoritmo proposto acima:

1. $\hat{T}_r$ está sendo usado no lugar de $\hat{T}$. Esta substituição é necessária uma vez que $\hat{T}$ é um número não necessariamente inteiro, e existe a necessidade de se ter um número inteiro $\hat{T}_r$ representando o número total estimado de tartarugas na população, número que é requerido para a construção de uma população finita

$$\{\hat{X}_j\}_{j=1}^{\hat{T}_r}.$$

2. A população $\{\hat{X}_j\}_{j=1}^{\hat{T}_r}$ de forma geral não segue exatamente a distribuição estimada $\hat{F}$, uma vez que uma população finita não pode sempre reproduzir exatamente uma distribuição de probabilidade arbitrária. Experimentos em

computador mostram, por este motivo, a necessidade de se calcular a correção *erro* no item 3 do algoritmo, utilizada no item 4.4 para a correção da média estimada pelo *bootstrap*. Sem esta correção haverá uma forte subcobertura no intervalo de confiança produzido pelo *bootstrap*.

As duas aproximações feitas tornam-se negligíveis à medida que $T \rightarrow \infty$.

8.3. Avaliação do *bootstrap*: metodologia de simulação

O intervalo de confiança produzido pelo algoritmo da Seção 8.1 teve suas propriedades estatísticas investigadas por meio de simulações em computador, feitas com base nas populações-modelo de tartarugas contidas em arquivos de dados construídos segundo metodologia descrita no Capítulo 3. Foram realizadas 160 simulações diferentes, correspondentes a 160 populações-modelo. Para $\omega_d = 5, 6, 7$ e 8 , para $T = 10, 20, 40, 100$ e 250 , foram utilizadas as distribuições AA, AY, BC, BM, FL, RI, SR e ST. As simulações, que incluíram os cálculos para a avaliação do teste do modelo (Capítulo 9), ocuparam no total cerca de 760 horas de tempo de processador no sistema IBM SP2.

Para cada população-modelo, a simulação tem por objetivo principal avaliar se a cobertura do intervalo de confiança produzido pelo *bootstrap* corresponde à cobertura nominal deste intervalo. Ou seja, para um α dado, $0 < \alpha < 0.5$, deseja-se investigar se $\Pr\{\hat{\mu}_{\alpha_1} \leq \mu \leq \hat{\mu}_{\alpha_2}\} = 1 - 2\alpha$. Não serão investigados os erros de cobertura lateral. A simulação para cada população-modelo foi feita de acordo com o algoritmo abaixo (Figura 8.2).

1. Definir um valor para ω_e , o número máximo de ninhos por tartaruga que será admitido nas estimações da simulação. Este ω_e é o parâmetro designado por ω no modelo proposto no Capítulo 5. Em todas as

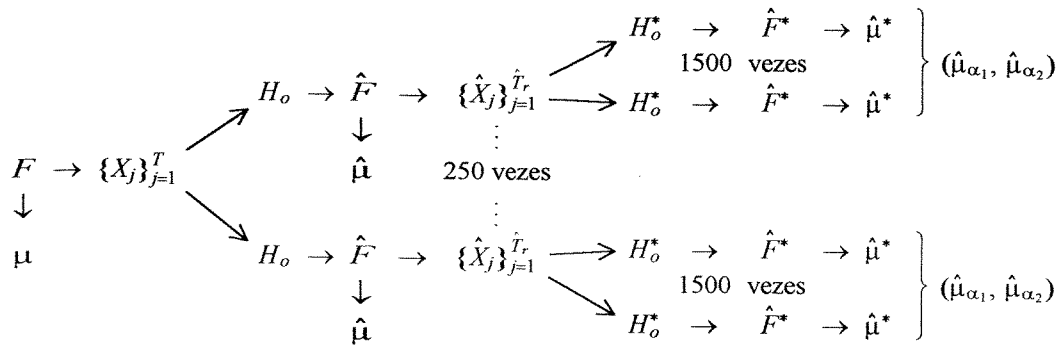


Figura 8.2. Simulação para avaliação do intervalo de confiança produzido pelo *bootstrap*.

simulações, $\omega_e = \omega_d$, onde ω_d é definido no Capítulo 3.

2. Conhecido, do arquivo de dados, o número T de tartarugas e $\{X_j\}_{j=1}^T$ (o número de ninhos de cada tartaruga da população), a distribuição F é calculada pelas equações (5A.1) e (5A.2), e a média μ desta distribuição F é calculada por (5A.3).

3. Para cada um dos seguintes valores de p : 0.98, 0.95, 0.90, 0.70, 0.50, 0.30, 0.20 e 0.15, executar os itens 3.1 - 3.2.

3.1. Repetir os itens 3.1.1 - 3.1.5 até que o item 3.1.5 seja executado 250 vezes:

3.1.1. Para $j = 1, \dots, T$, gerar $Y_j \sim \text{Bin}(X_j, p)$ (equação (5A.4)), onde X_j é o número de ninhos da tartaruga j e Y_j é o número de ninhos observados da tartaruga j .

3.1.2. Caso não exista nenhum $Y_j > 0$, $j = 1, \dots, T$, retornar ao item 3.1.1.

Se existir algum $Y_j > 0$, para as tartarugas com $Y_j > 0$ calcular

$H_o = (\gamma_1^o, \dots, \gamma_\omega^o)$, onde γ_k^o é computado pela equação (5.5).

3.1.3. Com $\hat{p}$ computado por (5.2), calcular $\hat{F}$ por (5.12), $\hat{\mu}$ por (5.13) e $\hat{T}$ por (5.1). No cálculo de $\hat{p}$, $N = \sum_j X_j$ e $N_o = \sum_{Y_j > 0} Y_j$.

3.1.4. Aplicar o algoritmo de construção do intervalo de confiança

(Seção 8.1) aos $\hat{F}$, $\hat{\mu}$ e $\hat{T}$ calculados no item 3.1.3, produzindo um intervalo de confiança $(\hat{\mu}_{\alpha_1}, \hat{\mu}_{\alpha_2})$. No item 5 do algoritmo da Seção 8.1, serão calculados percentis $\hat{\mu}_{\alpha_1}(\alpha)$ e $\hat{\mu}_{\alpha_2}(\alpha)$ relativos a $\alpha = 0.005$ (0.005) 0.25. Será assim produzido um conjunto de intervalos de confiança $(\hat{\mu}_{\alpha_1}(\alpha), \hat{\mu}_{\alpha_2}(\alpha))$ para a mesma estimativa $\hat{\mu}$, com coberturas nominais (para o algoritmo de Seção 8.1) respectivamente $1 - 2\alpha = 0.50$ (0.01) 0.99 (ao todo, são 50 intervalos de confiança diferentes). Estes diferentes intervalos de confiança para uma mesma estimativa $\hat{\mu}$ serão utilizados em um processo de calibração do intervalo produzido pelo *bootstrap* (Seção 8.4, abaixo).

3.1.5. Para cada valor de α , verificar se a desigualdade

$$\hat{\mu}_{\alpha_1}(\alpha) \leq \mu \leq \hat{\mu}_{\alpha_2}(\alpha) \text{ é verdadeira.}$$

3.2. Para cada valor de α , calcular a cobertura $c_{1-2\alpha}$ do intervalo

bootstrap, $c_{1-2\alpha} = \#\{\hat{\mu}_{\alpha_1}(\alpha) \leq \mu \leq \hat{\mu}_{\alpha_2}(\alpha) \text{ verdadeira}\} / 250$. $c_{1-2\alpha}$ é uma função de ω_e , T , d e p , sendo $d \in \{AA, AY, BC, BM, FL, RI, SR, ST\}$, o que será notado por $c_{1-2\alpha} = c_{1-2\alpha}(\omega_e, T, d, p)$.

8.4. Calibração do intervalo de confiança

A Figura 8.3, à esquerda, mostra os resultados das simulações para avaliação do *bootstrap* para $\omega_e = 7$ e $T = 100$, para uma cobertura nominal $1 - 2\alpha = 0.95$ (os resultados para outros valores de ω_e e T são semelhantes aos da Figura 8.3).

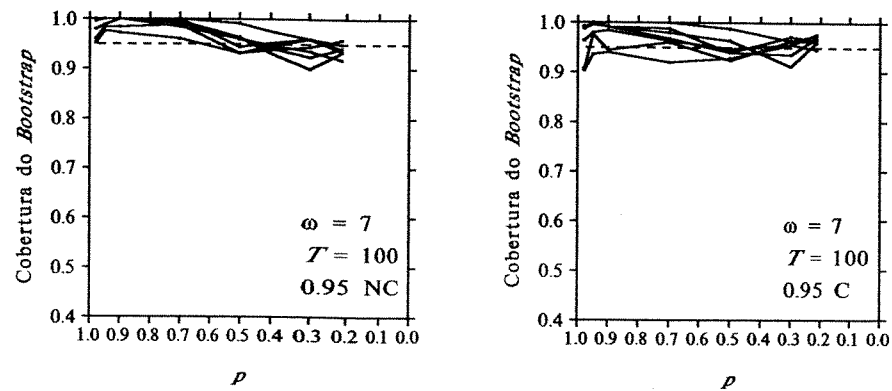


Figura 8.3. Cobertura do intervalo de confiança produzido pelo *bootstrap* em função de p , resultados obtidos por simulação para $\omega_e = 7$ e $T = 100$. Cada linha cheia representa uma população-modelo (veja o texto). Cobertura nominal a 0.95, indicada pela linha horizontal tracejada. Os resultados são apresentados apenas na região útil (aproximadamente). O gráfico da esquerda mostra $c_{0.95}$ para cada população-modelo, sem calibração. O gráfico da direita mostra os resultados das simulações após o processo de calibração descrito na Seção 8.4. Observar a escala reversa no eixo horizontal de cada gráfico.

Nesta figura, é plotado o valor $c_{0.95}$, em função de p , para cada população-modelo com $\omega_e = 7$ e $T = 100$, cada população sendo representada por uma linha cheia. Observa-se na figura que embora, de forma geral, para a cobertura nominal 0.95, o *bootstrap* esteja produzindo um intervalo de confiança com cobertura próxima à cobertura nominal, esta cobertura do intervalo produzido pelo *bootstrap* não é exata. Vê-se também na figura que esta cobertura depende, ainda que fracamente, de p . Estas observações sugerem (1) que, para se obter um intervalo de confiança com uma cobertura de 0.95, talvez se possa usar no processo de construção do intervalo pelo *bootstrap* um valor $1 - 2\alpha$ diferente de 0.95, de modo que a cobertura do intervalo produzido pelo *bootstrap* seja mais próxima de 0.95 do que aquela obtida com $1 - 2\alpha = 0.95$, e (2) que esta correção pode depender de p .

De acordo com esta idéia (ver Efron e Tibshirani, 1993, Capítulo 18), será feita uma calibração do intervalo de confiança produzido pelo *bootstrap*, utilizando-se a informação obtida com a utilização, nas simulações, de diferentes valores para α no item 3.1.4 do algoritmo da Seção 8.3. A calibração será feita de acordo com o seguinte algoritmo:

1. Fixar a cobertura desejada para o intervalo de confiança, que será a cobertura nominal. No presente trabalho, a cobertura nominal é igual a 0.95.
2. Para cada valor de ω_e , $\omega_e = 5, 6, 7$ e 8, para cada valor de T , $T = 10, 20, 40, 100$ e 250, para cada valor de p , $p = 0.98, 0.95, 0.90, 0.70, 0.50, 0.30, 0.20$ e 0.15, desde que $p \geq \bar{p}_{CV=0.25}(\omega_e, T)$ (ou seja, esteja dentro da região útil, ver Seção 6.4), calcular $k_{cal}(\omega_e, T, p)$ tal que

$$\sum_d (c_{k_{cal}(\omega_e, T, p)} - 0.95)^2 = \min_k \sum_d (c_k(\omega_e, T, d, p) - 0.95)^2, \quad (8.1)$$

onde $k = 0.50 (0.01) 0.99$ e $d = AA, AY, BC, BM, FL, RI, SR$ e ST (as várias distribuições-modelo), e onde $c_k(\omega_e, T, d, p)$, calculado no item 3.2 do algoritmo da Seção 8.3, é a cobertura produzida pelo *bootstrap* quando $1 - 2\alpha = k$, para o conjunto de parâmetros $\{\omega_e, T, d, p\}$. $k_{cal}(\omega_e, T, p)$ é a cobertura a ser utilizada no algoritmo de realização do *bootstrap* da Seção 8.1 (será usado neste algoritmo α tal que $1 - 2\alpha = k_{cal}(\omega_e, T, p)$), de modo a se ter um intervalo de confiança com a cobertura o mais próximo possível de 0.95, quando $p \in \{0.98, 0.95, \dots, 0.15\}$. A Figura 8.4 mostra os $k_{cal}(\omega_e, T, p)$ para $\omega_e = 5, 6, 7$ e 8. Nesta figura, os $k_{cal}(\omega_e, T, \bar{p}_{CV=0.25}(\omega_e, T))$ foram calculados por interpolação entre dois valores próximos de $k_{cal}(\omega_e, T, p)$, calculados por (8.1).

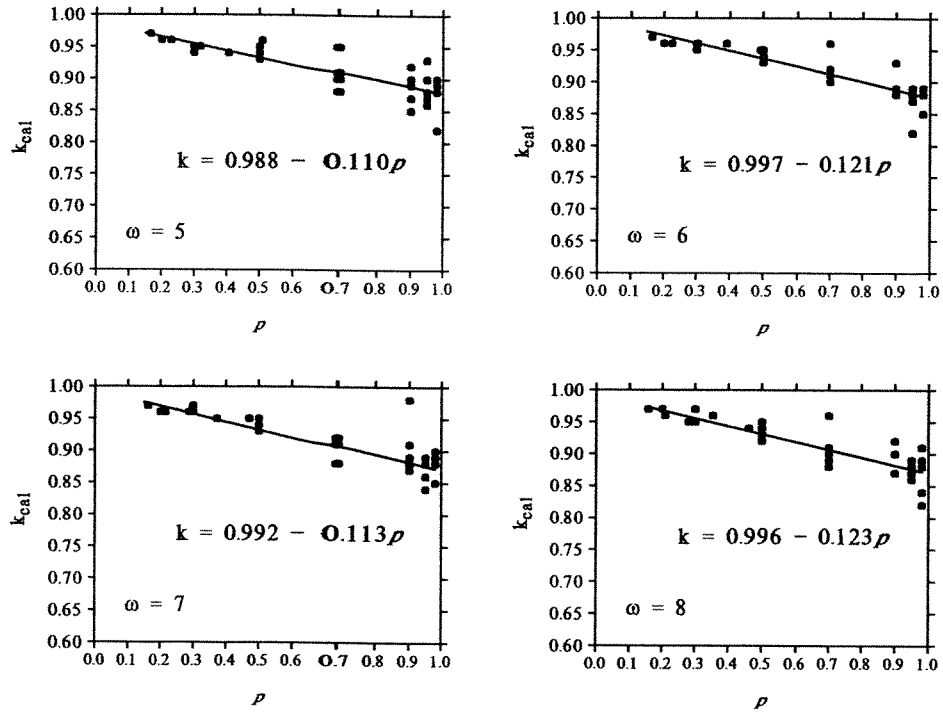


Figura 8.4. $k_{cal}(\omega_e, T, p)$, em função de p , para $\omega_e = 5, 6, 7$ e 8 . O gráfico para cada valor de ω_e mostra os k_{cal} para todos os valores de T e p com este ω_e . Em cada gráfico, a reta representa uma regressão linear entre os pontos plotados no gráfico, dando o valor da cobertura $1 - 2\alpha$ a ser utilizada no *bootstrap* (algoritmo da Seção 8.1) em função de p , $0.15 \leq p \leq 0.98$, para uma cobertura nominal de 0.95.

A Figura 8.3, à direita, mostra os resultados das simulações para $\omega_e = 7$ e $T = 100$, após a utilização do processo de calibração.

Na Figura 8.4, que mostra os $k_{cal}(\omega_e, T, p)$ calculados segundo o algoritmo descrito acima, estão representados apenas pontos referentes a $p = 0.98, 0.95, 0.90, 0.70, 0.50, 0.30, 0.20$ e 0.15 (e a alguns poucos valores interpolados). Estes valores foram os valores para p usados nas simulações para avaliação do *bootstrap*. No entanto, no uso real do *bootstrap* na estimação de μ , o valor $\hat{p}$ (item 1 do algoritmo da Seção 8.1) pode ser um valor qualquer, não

necessariamente igual a um dos valores de p utilizados nas simulações. Assim, para se usar a calibração do *bootstrap* com qualquer valor de $\hat{p}$, foi calculada para cada valor de ω_e , $\omega_e = 5, 6, 7$ e 8 , com base nos pontos apresentados em cada um dos gráficos da Figura 8.4, uma regressão linear da forma

$$k_{cal} = a + bp, \quad (8.2)$$

onde a e b são constantes. As equações obtidas estão apresentadas na Figura 8.4. Assim, dado um valor de $\hat{p}$ obtido pela equação (5.2), para o *bootstrap* será utilizado, no algoritmo da Seção 8.1, visando uma cobertura nominal igual a 0.95, um valor α tal que $1 - 2\alpha = k_{cal}(\hat{p})$, onde $k_{cal}(\hat{p})$ é obtido substituindo-se $\hat{p}$ na equação da forma (8.2) correspondente ao ω_e sendo utilizado na estimação, de acordo com a Figura 8.4.

8.5. Avaliação do *bootstrap*: resultados

A Figura 8.5 mostra os resultados obtidos em relação à cobertura do intervalo de confiança produzido pelo *bootstrap*, após a calibração, para uma cobertura nominal de 0.95, para $\omega_e = 7$. Nesta figura, cada linha cheia representa uma população-modelo, mostrando a cobertura do intervalo produzido pelo *bootstrap* para esta população-modelo em função de p . Os resultados para $\omega_e = 5, 6$ e 8 estão apresentados no Apêndice B.

8.6. Discussão

A Figura 8.5 mostra que, de forma geral, a cobertura do intervalo de confiança produzido pelo *bootstrap*, conforme os métodos propostos neste capítulo, é satisfatória. Em relação a um *bootstrap* por percentis (Efron e

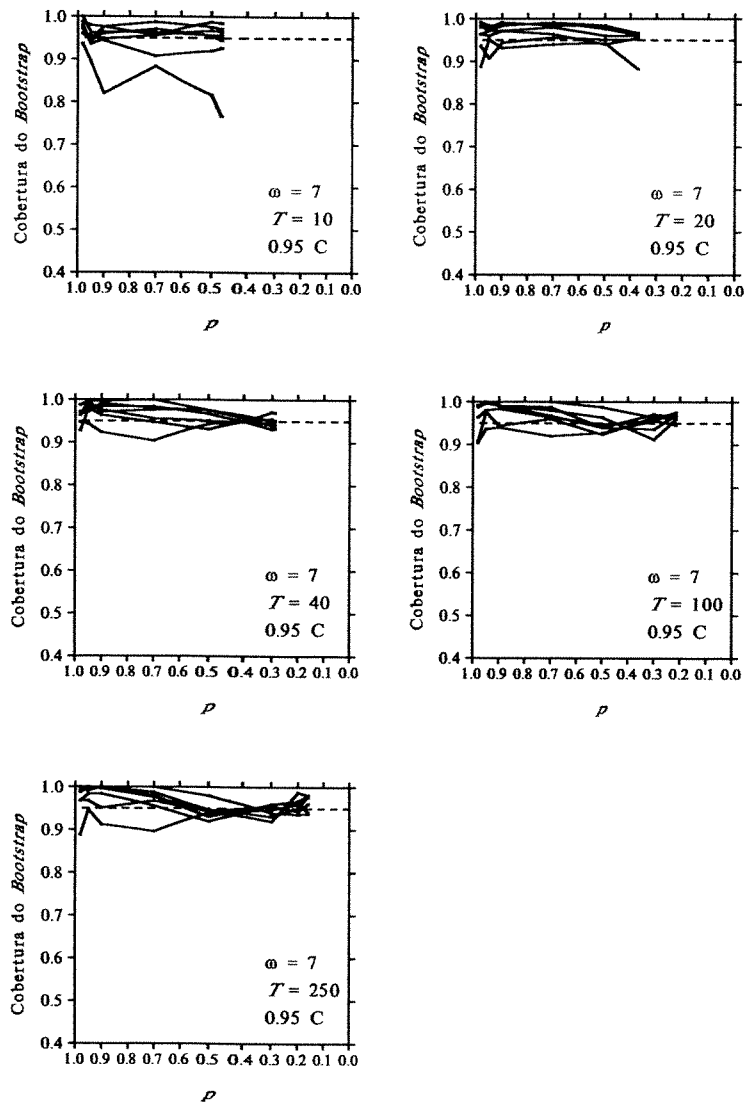


Figura 8.5. Cobertura do intervalo de confiança produzido pelo *bootstrap* em função de p , resultados obtidos por simulação para $\omega_e = 7$, com calibração. Cada linha cheia representa uma população-modelo (veja o texto). $T = 10, 20, 40, 100$ e 250 . A linha horizontal tracejada indica a cobertura nominal, igual a 0.95. Os resultados são apresentados apenas na região útil (aproximadamente) para cada T . Observar a escala reversa no eixo horizontal de cada gráfico.

Tibshirani, 1993, Capítulo 13), o método de *bootstrap* aqui proposto inclui três tipos de correção diferentes:

1. A correção *erro* (item 3 do algoritmo da Seção 8.1). Experimentos em computador mostram que, sem esta correção, existirá uma forte subcobertura no intervalo de confiança produzido pelo *bootstrap*. Esta correção é devida à diferença entre a distribuição dos $\{\hat{X}_j\}_{j=1}^{\hat{T}_r}$ e a distribuição $\hat{F}$ obtida no processo de estimação por (5.12).

2. O próprio *bootstrap* BCa é uma correção ao método básico do *bootstrap*, o *bootstrap* por percentis (Efron e Tibshirani, 1993, Capítulo 14). Experimentos em computador mostram que embora a diferença entre os resultados obtidos por um *bootstrap* por percentis e o *bootstrap* BCa não seja grande (ambos incluindo a correção do item 1), esta diferença é nítida o bastante para justificar o uso de um *bootstrap* mais complexo como é o *bootstrap* BCa.

3. A correção final é o processo de calibração. A Figura 8.3 mostra que há, com a calibração, uma melhora na cobertura do intervalo produzido pelo *bootstrap*, embora a calibração, ao menos na forma como está sendo feita, não seja um remédio para todas as imperfeições do *bootstrap*.

Experimentos em computador indicam que a correção do item 1 é a mais importante. Resultados razoáveis seriam obtidos apenas com esta correção e um *bootstrap* por percentis, sem calibração.

A Figura 8.5 (veja também o Apêndice B) mostra que para $T = 10$ há uma distribuição para a qual o *bootstrap* não está conseguindo produzir um intervalo de confiança com cobertura próxima à cobertura nominal (no caso, igual a 0.95). Nesta figura, a linha com cobertura nitidamente menor do que as demais linhas corresponde à população-modelo referente à distribuição AA. Esta distribuição

(Figura 3.1) tem uma forte assimetria negativa, o que a aproxima da distribuição de probabilidade para o número de ninhos por tartaruga com $p(1) = 1.0$, que é a distribuição que admite apenas a possibilidade de haver 1 ninho por tartaruga.

Vamos denotar esta distribuição por F_1 .

Para a distribuição F_1 , temos que $f(F_1) = F_1$ (veja o Capítulo 5), o que implica que, se $H_o = F_1$, então, qualquer que seja o valor de $\hat{p}$, $\hat{F} = F_1$. Ou seja, se a distribuição observada H_o é igual a F_1 , o processo de estimação definido no Capítulo 5 não tem como recuperar (de forma aproximada) a distribuição F que gerou $H_o = F_1$, caso F seja diferente de F_1 . Neste caso, o *bootstrap*, que depende de $\hat{F}$, produzirá um intervalo de confiança com largura igual a zero, ou seja, a estimativa $\hat{\mu}$ será igual a 1 com erro igual a zero.

Com $T = 10$, um número pequeno de tartarugas, para uma distribuição como a distribuição AA, com forte assimetria negativa, a distribuição H_o será, com alta probabilidade, tanto maior quanto menor o valor de p , igual a F_1 , o que fará com que o intervalo de confiança produzido pelo *bootstrap* não inclua a média real μ , o que aparecerá nas simulações para avaliação do *bootstrap* (Seção 8.3) como um valor $c_{1-2\alpha}$ (item 3.2 do algoritmo da Seção 8.3) com valor bem menor que a cobertura nominal.

Em outros termos, existe no α -espaço uma vizinhança da distribuição F_1 tal que, se a distribuição real F pertence a esta vizinhança, $F \neq F_1$, o processo de estimação terá dificuldade em recuperar (de forma aproximada) esta distribuição real F , pois com alta probabilidade o processo de observação produzirá $H_o = F_1$. Esta dificuldade, para uma F fixa, diminuirá à medida que T aumentar e/ou que p aumentar, pois neste caso, se $F \neq F_1$, será menor a probabilidade de que $H_o = F_1$.

CAPÍTULO 9

Teste do modelo

Um teste de ajuste do modelo do Capítulo 5 será feito por meio de um teste de distância mínima estocástico (Beran, 1984; Beran e Millar, 1987), em que a distribuição nula é estimada através de um *bootstrap*. O teste é naturalmente relacionado ao método definido no Capítulo 5, onde a estimação de μ é feita por um processo de distância mínima.

9.1. O teste

A estatística t será definida por

$$t = \left\| H_o - \hat{f}(\hat{F}_{\hat{X}}) \right\|, \quad (9.1)$$

onde H_o é a distribuição observada de ninhos por tartaruga (Capítulo 5), $\hat{F}_{\hat{X}}$ é a distribuição de probabilidade de $\{\hat{X}_j\}_{j=1}^{\hat{T}_r}$ (item 2 do algoritmo da Seção 8.1), $\hat{f}$ é a função f definida por (5.9) - (5.10) computada com $\hat{p}$ (definido por (5.2)) no lugar de p e $\|\cdot\|$ é a norma euclideana em R^ω , sendo ω o valor máximo para o número real de ninhos por tartaruga admitido na estimação. A distribuição $\hat{F}_{\hat{X}}$ é computada por

$$\hat{F}_{\hat{X}} = (\delta_1, \delta_2, \dots, \delta_\omega), \quad (9.2)$$

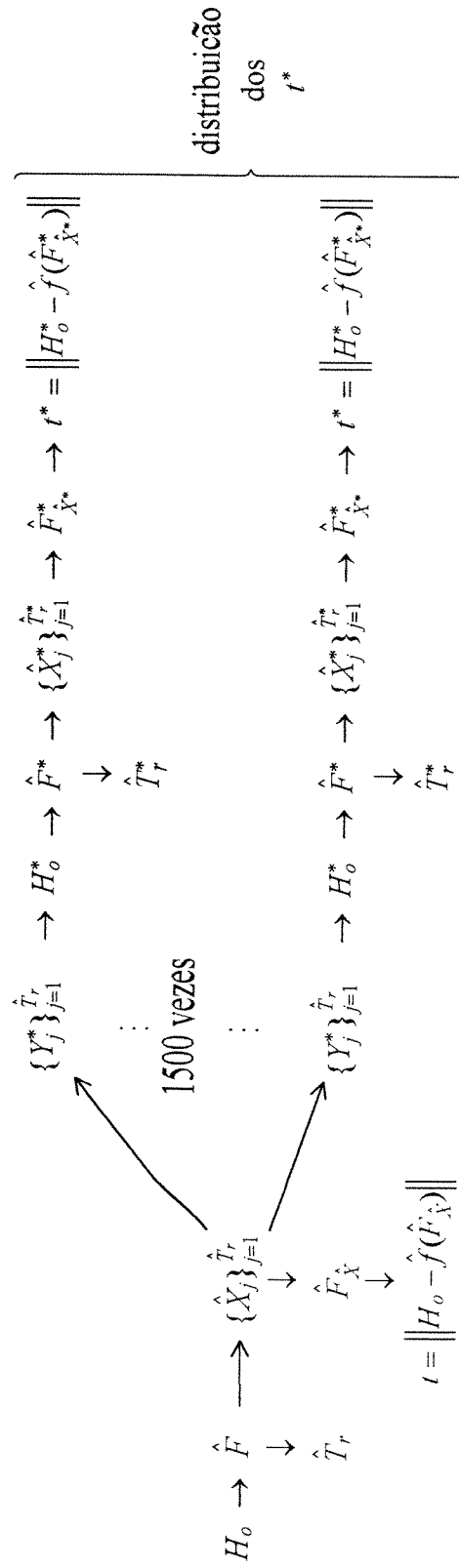


Figura 9.1. Teste do modelo: *bootstrap* para estimação da distribuição nula da estatística t . O valor t será comparado com um valor crítico obtido da distribuição dos t^* .

onde

$$\delta_k = \# \{ \hat{X}_j \mid \hat{X}_j = k \} / \hat{T}_r, \quad k = 1, \dots, \omega. \quad (9.3)$$

A estatística t será comparada com um valor crítico obtido de uma distribuição nula estimada por um *bootstrap*. Em termos práticos, é claro, esta distribuição será estimada dentro do mesmo *bootstrap* que é utilizado na construção do intervalo de confiança (Capítulo 8). O algoritmo para a construção da distribuição nula (Figura 9.1) será descrito como uma extensão do algoritmo da Seção 8.1 (designado por ALG 8.1), que descreve o *bootstrap* utilizado na construção do intervalo de confiança:

1. Definir um valor α tal que, se $t_{1-\alpha}$ é o percentil $1 - \alpha$ da distribuição dos t^* (calculados abaixo), o teste rejeitará a hipótese de que o modelo correto é definido pela distribuição $\hat{F}$ (calculada no Capítulo 5) quando $t > t_{1-\alpha}$. Um valor usual é $\alpha = 0.05$ (o erro de tipo I admitido no teste).

2. Dentro do item 4 de ALG 8.1, acrescentar os itens 4.5 - 4.8, definidos a seguir:

4.5. Calcular $\hat{T}_r^*$, aplicando (5.1) a $\hat{\mu}^*$ (item 4.4 de ALG 8.1) e arredondando o resultado para o inteiro mais próximo.

4.6. Construir $\{\hat{X}_j^*\}_{j=1}^{\hat{T}_r^*}$ seguindo da forma mais próxima possível a distribuição $\hat{F}^*$ calculada no item 4.3 de ALG 8.1. Esta construção é análoga à construção do item 2 de ALG 8.1.

4.7. Calcular $\hat{F}_{\hat{X}^*}^*$, a distribuição de probabilidade de $\{\hat{X}_j^*\}_{j=1}^{\hat{T}_r^*}$, definida por

$$\hat{F}_{\hat{X}^*}^* = (\delta_1^*, \delta_2^*, \dots, \delta_\omega^*), \quad (9.4)$$

onde

$$\delta_k^* = \# \{ \hat{X}_j^* \mid \hat{X}_j^* = k \} / \hat{T}_r^*, \quad k = 1, \dots, \omega. \quad (9.5)$$

4.8. Calcular $t^* = \left\| H_o^* - \hat{f}(\hat{F}_{\hat{X}^*}^*) \right\|$, onde $\hat{f}$ e $\|\cdot\|$ são os mesmos do cálculo da equação (9.1) e H_o^* foi calculado no item 4.2 de ALG 8.1.

No item 4 de ALG 8.1, então, serão agora executados repetidas vezes os itens 4.1 - 4.8.

3. Da distribuição dos B valores t^* calculados no item 4.8 de ALG 8.1, calcular $t_{1-\alpha}$, o percentil $1 - \alpha$ desta distribuição, que será o valor crítico utilizado no teste na comparação com a estatística t .

9.2. Avaliação do teste: metodologia de simulação

A avaliação do teste proposto na seção anterior foi feita dentro das simulações para avaliação do *bootstrap* (Seção 8.3), com $B = 1500$. O algoritmo para a avaliação do teste será apresentado como uma extensão do algoritmo da Seção 8.3 (designado por ALG 8.3):

1. No item 3.1.3 de ALG 8.3, calcular $\hat{F}_{\hat{X}^*}$, definida por (9.2) - (9.3), e t , definido por (9.1).
2. No lugar do item 3.1.4 de ALG 8.3, aplicar o algoritmo da Seção 9.1, produzindo um valor crítico $t_{1-\alpha}$. Este cálculo será feito para $\alpha = 0.01$ (0.01) 0.50, visando um processo de calibração do teste do modelo, de forma análoga ao que foi feito no Capítulo 8 em relação ao *bootstrap*.
3. No lugar do item 3.1.5 de ALG 8.3, para cada valor de α , verificar se a desigualdade $t \leq t_{1-\alpha}$ é verdadeira.
4. No lugar do item 3.2 de ALG 8.3, para cada valor de α (item 2 acima), calcular $\pi_{1-\alpha} = \#\{t \leq t_{1-\alpha} \text{ verdadeira}\} / 250$. $\pi_{1-\alpha}$ é uma função de ω_e , T , d e p , o que será notado por $\pi_{1-\alpha} = \pi_{1-\alpha}(\omega_e, T, d, p)$, onde $d \in \{AA, AY,$

BC, BM, FL, RI, SR, ST}. $\pi_{1-\alpha}$ mede, em termos de probabilidade, o tamanho da região de aceitação do teste para o valor crítico $t_{1-\alpha}$.

9.3. Calibração do teste

O teste será calibrado por um processo análogo ao da Seção 8.4, onde é feita uma calibração do intervalo de confiança produzido pelo *bootstrap*. O algoritmo para a calibração é o seguinte:

1. Fixar o valor de α que se deseja no teste. No presente trabalho, $\alpha = 0.05$.
2. Para cada valor de ω_e , $\omega_e = 5, 6, 7$ e 8 , para cada valor de T , para cada valor e p , $p = 0.98, 0.95, 0.90, 0.70, 0.50, 0.30, 0.20$ e 0.15 , desde que p esteja na região útil (Seção 6.4), calcular $\pi_{cal}(\omega_e, T, p)$ tal que

$$\sum_d (\pi_{cal}(\omega_e, T, p) - 0.05)^2 = \min_k \sum_d (\pi_k(\omega_e, T, d, p) - 0.05)^2, \quad (9.6)$$

onde $k = 0.50 (0.01) 0.99$, e $d = AA, AY, BC, BM, FL, RI, SR$ e ST , e onde $\pi_k(\omega_e, T, d, p)$ é calculado no item 4 do algoritmo da Seção 9.2. A Figura 9.2 mostra os $\pi_{cal}(\omega_e, T, p)$ para $\omega_e = 5, 6, 7$ e 8 . Nesta figura, os $\pi_{cal}(\omega_e, T, \bar{p}_{CV=0.25}(\omega_e, T))$ foram calculados por interpolação entre dois valores próximos de $\pi_{cal}(\omega_e, T, p)$, calculados por (9.6).

Do mesmo modo que no Capítulo 8, foi feita, para cada valor de ω_e , $\omega_e = 5, 6, 7$ e 8 , uma regressão linear entre os pontos da Figura 9.2 referentes a cada ω_e . As equações de regressão, da forma

$$\pi_{cal} = a + bp, \quad (9.7)$$

onde a e b são constantes, são mostradas na Figura 9.2. Por meio destas equações, dado um valor $\hat{p}$ calculado por (5.2), será calculado o valor

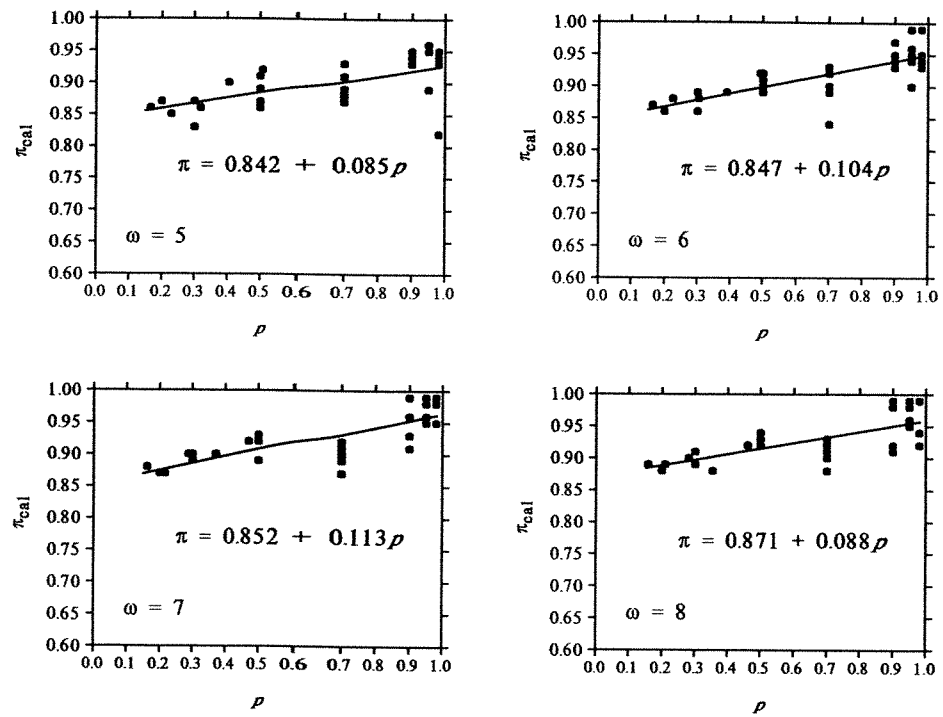


Figura 9.2. $\pi_{cal}(\omega_e, T, p)$, em função de p , para $\omega_e = 5, 6, 7$ e 8 . O gráfico para cada valor de ω_e mostra os π_{cal} para todos os valores de T e p com este ω_e . Em cada gráfico, a reta representa uma regressão linear entre os pontos plotados no gráfico, dando, em função de p , o valor $1 - \alpha$ tal que α é o erro de tipo I a ser utilizado no item 1 do algoritmo da Seção 9.1, para se obter um teste com erro de tipo I real igual (aproximadamente) a 0.05.

$\pi_{cal}(\hat{p}) = a + b\hat{p}$, de modo que $\alpha_{cal}(\hat{p}) = 1 - \pi_{cal}(\hat{p})$ será o valor de α a ser utilizado no item 1 do algoritmo da Seção 9.1, de modo a ser obter um teste com um erro de tipo I o mais próximo possível de 0.05.

9.4. Avaliação do teste: resultados

A Figura 9.3 mostra os resultados obtidos com as simulações do teste, após a calibração, para um erro $\alpha = 0.05$ e $\omega_e = 7$. Nesta figura, cada linha cheia representa um população-modelo, mostrando a probabilidade de aceitação do

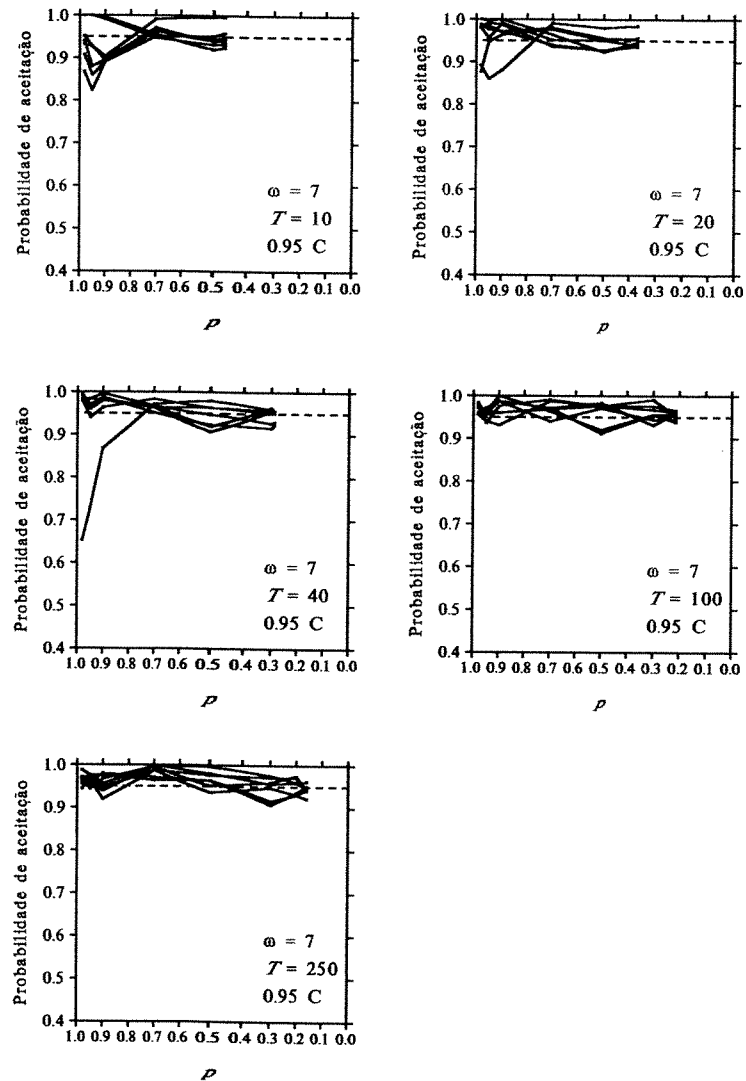


Figura 9.3. Probabilidade de aceitação do teste, em função de p , resultados obtidos por simulação para $\omega_e = 7$, com calibração. Cada linha cheia representa uma população-modelo (veja o texto). $T = 10, 20, 40, 100$ e 250 . A linha horizontal tracejada indica $1 - \alpha = 0.95$, onde $\alpha = 0.05$ é o erro de tipo I desejado para o teste. Os resultados são apresentados apenas na região útil (aproximadamente) para cada T . Observar a escala reversa no eixo horizontal de cada gráfico.

teste para esta população-modelo em função de p . Os resultados para $\omega_e = 5, 6$ e 8 estão apresentados no Apêndice C.

9.5. Discussão

O teste, na forma como é construído, conforme mostra a Figura 9.3 (veja também o Apêndice C), apresenta resultados razoáveis. Para $\omega_e = 7$ e 8 , para $T = 40$ principalmente, mas também, de forma menos forte, para $T = 10$, para valores altos de p , a região de aceitação é menor do que deveria para distribuições com assimetria negativa forte, como a distribuição AA. O teste aparentemente é mais sensível a uma assimetria negativa forte do que o método de construção do intervalo de confiança, uma vez que um exame dos arquivos de dados produzidos pelas simulações mostra que mesmo distribuições nem tão negativamente assimétricas, como as distribuições RI e SR, produzem resultados às vezes aquém dos que seriam desejados. A este respeito, veja a discussão no final do Capítulo 8.

Um resultado não-significante no teste deste capítulo não significa que o modelo seja necessariamente correto. Uma análise da potência do teste (ou, equivalentemente, do erro de tipo II), o que acarretaria toda uma outra série de simulações, seria necessária para a avaliação de forma completa dos resultados do teste.

CAPÍTULO 10

Exemplo: aplicação do modelo a dados das tartarugas marinhas *Caretta caretta* da Praia do Forte, Bahia

Como exemplo de aplicação do modelo desenvolvido neste trabalho, foram analisados dados colhidos pelo Projeto TAMAR na Praia do Forte, estado da Bahia, referentes a desovas das tartarugas *Caretta caretta*.

10.1. Metodologia de campo do Projeto TAMAR

A Praia do Forte (12°34'S, 37°58'W), no estado da Bahia, está localizada a cerca de 80 km ao norte de Salvador, no município de Mata de São João, e tem cerca de 13 km de comprimento, sendo limitada ao sul pelo rio Pojuca e ao norte pelo rio Imbassaí. A região onde está situada esta praia, o litoral norte do estado da Bahia, é uma das principais áreas de desova das tartarugas marinhas no Brasil. Na Praia do Forte desovam quatro das cinco espécies de tartarugas marinhas encontradas em território brasileiro: *Caretta caretta*, *Eretmochelys imbricata*, *Chelonia mydas* e *Lepidochelys olivacea*. Nesta praia, sede de sua coordenação nacional, o Projeto TAMAR exerce um trabalho de conservação das tartarugas marinhas que inclui, além da proteção a ninhos e tartarugas, atividades de educação ambiental junto a pescadores, escolas, turistas e à comunidade da área em geral.

Na Praia do Forte, a temporada de desova é geralmente o período indo de setembro a março. Uma temporada de desova é designada nos arquivos do Projeto TAMAR pelo termo *campanha*, e registrada com um código que inclui o ano inicial e o ano final da temporada. Assim, por exemplo, a temporada de desova indo de setembro de 1987 a março de 1988 é registrada como sendo a campanha 87/88. Das desovas registradas pelo Projeto TAMAR na Praia do Forte desde a campanha 87/88 até a campanha 94/95, 83% são da espécie *Caretta caretta*, 15% são de *Eretmochelys imbricata*, 1% são de *Chelonia mydas* e 1% de *Lepidochelys olivacea*. Para as tartarugas *Caretta caretta*, cerca de 93% das desovas são feitas nos meses de outubro, novembro, dezembro e janeiro, com o pico do número de desovas ocorrendo no mês de novembro.

A metodologia de campo do Projeto TAMAR inclui o monitoramento da praia à noite e o monitoramento da praia pela manhã. Além do trabalho na praia, existe ainda o trabalho relativo ao cercado de incubação, que é um trecho de areia situado na parte alta da praia, delimitado e protegido (como indica o nome) por uma cerca, exposto a sol e chuva plenos, para onde são transferidas parte das desovas classificadas como de risco pela sua situação na praia. A metodologia utilizada no trato do cercado de incubação não será apresentada em detalhe.

1. Monitoramento da praia à noite - marcação das fêmeas. A cada noite, os 13 km da praia são percorridos por uma equipe em um jipe com tração nas quatro rodas, durante um número variável de horas. Toda ocorrência é registrada em uma ficha de campo, cujos dados são posteriormente transferidos para um arquivo em computador. Por *ocorrência* entende-se o encontro de uma tartaruga na praia, o encontro de um ninho, mesmo que a tartaruga não tenha sido encontrada por ocasião da desova, ou o encontro de um rastro, mesmo sem haver

um ninho a ele associado. Cada ocorrência tem a data, hora, local (trecho da praia) e tipo da ocorrência registrados.

No caso de encontro de uma fêmea na praia, a tartaruga é medida, identificada quanto à espécie e lhe são aplicadas marcas numeradas. Na fase inicial do Projeto TAMAR na Praia do Forte era colocada uma única marca em cada tartaruga, mas atualmente, com a finalidade de minimizar o problema da perda de marcas (veja o Capítulo 7), utiliza-se o processo de marcação dupla (Limpus, 1992b), sendo colocada uma marca em cada uma das nadadeiras dianteiras. Caso a tartaruga ao vir do mar já porte marcas, estas são registradas e substituídas, caso não estejam em bom estado.

Quando, sem que a tartaruga seja observada, é encontrado um rastro que inclui uma cama, havendo toda a aparência de ter ocorrido uma desova, há o problema de se confirmar ou não a existência do ninho. Como a tartaruga pode ter realizado todo o processo de desova, com a confecção da cama, e não ter de fato desovado, a existência da cama e de uma grande área da praia com a areia revirada não é garantia de haver um ninho. Caso o ninho seja encontrado, a ocorrência é classificada como CD (com desova). Caso o ninho não seja encontrado pela equipe da noite, haverá ainda na manhã seguinte uma tentativa de encontrá-lo. Caso o ninho de fato não seja encontrado, a ocorrência é classificada como SD (sem desova).

Caso seja realizada uma desova, duas coisas podem ocorrer: (1) se houver necessidade de transferir o ninho de local, o que pode acontecer porque a tartaruga colocou o ninho em um local naturalmente de risco ou em um local sujeito a perturbações pela atividade humana, os ovos, segundo técnica padrão (Pritchard *et al.*, 1983), são colocados em uma caixa de isopor forrada com areia

e transferidos para um ninho artificialmente construído em outro local. Este outro local pode ser na própria praia, onde, após a transferência, a incubação segue sob condições naturais, ou no cercado de incubação, com a incubação ocorrendo sob o cuidado de técnicas de manejo. (2) Caso o ninho depositado pela tartaruga não esteja em risco aparente e não esteja na área definida *a priori* como de transferência (o trecho da praia próximo ao vilarejo, com grande movimentação de pessoas), este ninho é deixado *in situ*, incubando sob condições naturais.

Os ninhos *in situ* são abertos após a eclosão para a obtenção de uma contagem do número de ovos, realizada através da contagem das cascas, e, no caso da tartaruga que colocou o ninho não ter sido encontrada no momento da desova, para a determinação da espécie, o que é feito examinando-se filhotes natimortos ou embriões grandes que quase sempre existem no ninho. No caso dos ninhos no cercado de incubação, a determinação da espécie é feita através dos filhotes, que ao emergirem do ninho ficam retidos em uma tela colocada ao seu redor e são posteriormente liberados na praia.

2. Monitoramento da praia pela manhã. A cada manhã, a praia é percorrida a pé por dois pescadores contratados pelo Projeto TAMAR, cada um deles cobrindo cerca de metade da extensão da praia. Estes pescadores são homens experientes no reconhecimento de rastros e ninhos de tartarugas marinhas, alguns deles sendo antigos tartarugueiros, que no passado coletavam ovos das tartarugas e freqüentemente matavam as próprias fêmeas. O trabalho destes pescadores é registrar todas as desovas ocorridas na noite imediatamente anterior, já que, como o jipe não fica a noite inteira na praia, muitas desovas ocorrem sem serem detectadas pela equipe da noite. Outra função destes

pescadores é confirmar ou não a existência de um ninho que eventualmente não tenha sido encontrado pela equipe da noite ao ter sido achado um rastro com cama associada. Todo o trabalho pela manhã é supervisionado pela equipe técnica do Projeto TAMAR, que é responsável pela execução das atividades de campo, e os dados obtidos são também passados para computador. No caso da Praia do Forte, em função deste trabalho feito a pé a cada manhã ao longo de toda a extensão da praia, pode-se considerar que todos os ninhos depositados na praia são localizados pelo Projeto TAMAR.

10.2. Metodologia aplicada aos dados da Praia do Forte

Desde a campanha 82/83, campanha inicial do Projeto TAMAR na Praia do Forte, até a campanha 94/95, um total de 1301 ocorrências identificadas à espécie *Caretta caretta* foram registradas naquela praia. Estes 1301 registros estavam classificados, quanto ao tipo da ocorrência, nas seguintes categorias: com desova (código CD), sem desova (código SD), meia-lua (código ML) e desova perdida ou predada (código DP). O código DP representa um ninho destruído por erosão da praia, atacado por predadores ou perdido total ou parcialmente por qualquer outro motivo.

Os 1301 registros foram processados da forma exposta abaixo. O objetivo principal deste processamento é agrupar os dados por tartaruga individual e conhecer o número de ninhos observados para cada tartaruga a partir dos dados de campo. (Para a definição de "ninho observado" e "tartaruga observada", veja o Capítulo 4.)

1. O arquivo de dados do Projeto TAMAR é um arquivo de tipo .DBF (padrão Dbase). Este arquivo foi processado por um programa escrito na

linguagem CA-Clipper 5.2, programa que agrupa as tartarugas marcadas, formando blocos de registros para cada tartaruga individual. Cada bloco de registros, correspondente a uma tartaruga, tem a forma do exemplo a seguir, referente a uma das tartarugas *Caretta caretta* encontradas na Praia do Forte:

N_Arq	Campanha	Data	Praia	Tipo	Espécie	Marcas
4100	87/88	20/11/87	PF	CD	CC	3010 3011
4120	87/88	05/12/87	PF	CD	CC	3010 3011
9813	90/91	09/11/90	PF	SD	CC	3010 3011 5040
9815	90/91	10/11/90	PF	ML	CC	3010 3011 5040

No exemplo acima, a coluna "Tipo" corresponde ao tipo da ocorrência, e N_Arq é o número do registro no arquivo de dados. No total, foram criados 133 blocos de registros, ou seja, da campanha 82/83 até a campanha 94/95 um total de 133 tartarugas *Caretta caretta* diferentes foram observadas na Praia do Forte. Estes 133 blocos correspondem a 269 registros diferentes, ou seja, da campanha 82/83 até a campanha 94/95 houve na Praia do Forte 269 episódios de marcação de tartarugas *Caretta caretta*. Note-se que "marcação" aqui significa colocação de marcas novas ou simples observação de marcas já existentes.

2. Dos 269 registros, 11 foram eliminados, por serem observações correspondentes a um episódio de reprodução já representado em outro registro da mesma tartaruga. Por exemplo, se uma tartaruga tem um registro com código SD e outro registro com código CD apenas 1 dia após o registro SD, o registro SD será eliminado, pois os dois registros correspondem a uma mesma desova, sendo que na noite anterior à desova a tartaruga veio à praia mas não desovou. No exemplo anterior, o registro com número no arquivo 9813 foi eliminado, e o bloco de dados daquela tartaruga fica agora assim:

N_Arq	Campanha	Data	Praia	Tipo	Espécie	Marcas		
4100	87/88	20/11/87	PF	CD	CC	3010	3011	
4120	87/88	05/12/87	PF	CD	CC	3010	3011	
9815	90/91	10/11/90	PF	ML	CC	3010	3011	5040

Note-se no exemplo acima que estamos inferindo a existência de uma desova em data bem próxima a 10/11/1990, que foi a data em que uma meia-lua (ML) foi registrada para esta tartaruga. A tartaruga não foi observada realizando esta desova. O ninho, no entanto, foi certamente computado no total de ninhos para a campanha 90/91. É um dos muitos ninhos desta campanha de que não se sabe a que tartaruga pertencem. O objetivo de utilizar desta forma registros com código SD ou ML é aproveitar ao máximo as informações disponíveis no arquivo de dados.

Com as 11 eliminações, o arquivo com os 133 blocos de registros passa a ter um total de 258 registros. Este número, 258, será tomado como sendo o número de ninhos observados entre as campanhas 82/83 e 94/95.

3. Paralelamente ao processamento acima, um programa escrito em CA-Clipper 5.2 calculou, a partir do arquivo de dados original com 1301 registros, o número de ninhos para cada campanha entre 82/83 e 94/95. Foram computadas como ninhos os registros com tipo de ocorrência com código CD ou DP. Foi criado um arquivo com o número de ninhos a cada campanha entre 82/83 e 94/95.

4. O arquivo com os 258 registros agrupados por tartaruga e o arquivo com o número de ninhos por campanha foram processados por um programa escrito na linguagem Turbo Pascal 7.0. Este programa calculou:

4.1. Para cada campanha, a probabilidade estimada de que um ninho seja

observado pela equipe de campo. Para cada campanha, esta probabilidade é definida por $\hat{p} = \text{número de ninhos observados na campanha} / \text{número de ninhos na campanha}$.

4.2. Para cada campanha, a distribuição do número de tartarugas segundo o número de ninhos observados.

4.3. Para cada campanha, o número médio observado de desovas por tartaruga, μ_o . Esta média é definida por $\mu_o = \text{número total de ninhos observados na campanha} / \text{número de tartarugas diferentes (indivíduos) observadas na campanha}$.

10.3. Os dados da Praia do Forte: resultados e discussão

A Figura 10.1 mostra o número de ninhos entre as campanhas 82/83 e 94/95.

A série temporal do número de desovas apresenta um salto nítido entre as

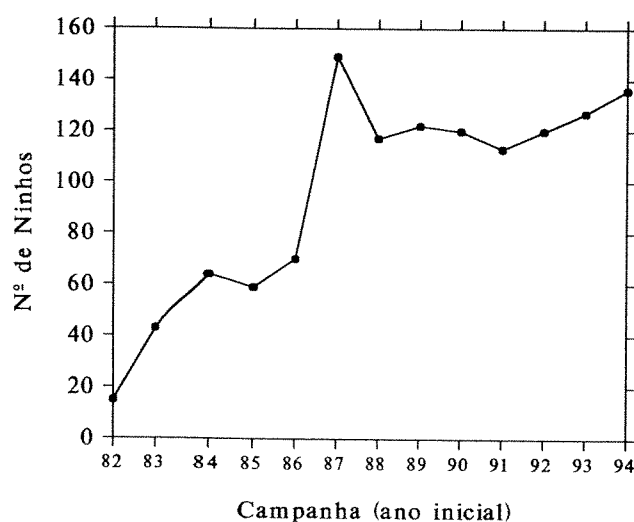


Figura 10.1. Número de ninhos por campanha, Praia do Forte, BA, espécie *Caretta caretta*, campanhas 82/83 a 94/95.

campanhas 86/87 e 87/88. Uma análise da distribuição temporal dos registros (não apresentada) indica que até a campanha 86/87 não houve uma cobertura completa da temporada de desova pela equipe de campo do Projeto TAMAR. Por este motivo, serão utilizados a seguir apenas os dados da campanha 87/88 em diante.

A Tabela 10.1 mostra o número de ninhos e de ninhos observados desde a campanha 87/88, bem como, para cada campanha, a probabilidade estimada de observação de cada ninho, $\hat{p}$. Observa-se que esta probabilidade varia conforme a campanha, indicando um esforço variável no trabalho de campo a cada campanha.

A Tabela 10.2 mostra a distribuição do número de tartarugas segundo o número de ninhos observados e o número médio de ninhos observados por tartaruga. O número máximo de ninhos observados para uma mesma tartaruga é de 5 ninhos, alcançado nas temporadas 87/88 e 89/90.

Tabela 10.1. Total de ninhos, ninhos observados e probabilidade estimada de observação de cada ninho por campanha, Praia do Forte, BA, espécie *Caretta caretta*, campanhas 87/88 a 94/95

campanha	ninhos observados	total de ninhos	probabilidade de observação
87/88	51	149	0.3423
88/89	23	117	0.1966
89/90	29	122	0.2377
90/91	42	120	0.3500
91/92	12	113	0.1062
92/93	12	120	0.1000
93/94	20	127	0.1575
94/95	18	136	0.1324

Tabela 10.2. Distribuição do número de tartarugas segundo o número de ninhos observados por tartaruga e número médio de ninhos observados por tartaruga por campanha, Praia do Forte, BA, espécie *Caretta caretta*, campanhas 87/88 a 94/95

campanha	nº de tartarugas segundo o nº de ninhos observados					nº médio de ninhos observados por tartaruga
	1	2	3	4	5	
87/88	12	9	4	1	1	1.89
88/89	15	1	2	-	-	1.28
89/90	15	3	1	0	1	1.45
90/91	22	5	2	1	-	1.40
91/92	10	1	-	-	-	1.09
92/93	12	-	-	-	-	1.00
93/94	7	2	3	-	-	1.67
94/95	16	1	-	-	-	1.06

A eliminação dos 11 registros correspondentes a episódios de reprodução já representados em outros registros da mesma tartaruga é fundamentada no reconhecimento pelos pesquisadores da biologia de reprodução das tartarugas *Caretta caretta* (Richardson, 1982; Limpus, 1985) de que uma fêmea que venha à praia e não desove retornará à praia dentro de poucos dias, freqüentemente na mesma noite ou na noite seguinte, para realizar outra tentativa de desova. O intervalo de tempo entre duas tentativas é consideravelmente menor que o período internidal normal entre duas desovas completas. Richardson (1982) reporta que, para uma amostra de 589 *Caretta caretta* observadas na Georgia, USA, entre 1964 e 1977, nenhuma retornou para uma nova desova em menos de 9 dias se a desova inicial foi completa, enquanto que desovas não realizadas ou interrompidas eram novamente tentadas ou completadas usualmente em um ou

dois dias. Limpus (1985), estudando as *Caretta caretta* na Austrália entre 1976/1977 e 1979/1980, reporta que 80% das tartarugas que tiveram uma tentativa não realizada de desova foram observadas retornando para nova tentativa em no máximo 5 dias, com intervalo médio de 1.08 dias. Este mesmo pesquisador reporta o período internidal mínimo entre desovas completas como sendo de 9 dias.

Pode-se inferir que a tartaruga que na campanha 87/88 da Praia do Forte teve 5 ninhos observados desovou na realidade no mínimo 6 vezes, se aceitarmos que esta tartaruga realizou uma desova não observada entre duas desovas observadas distantes entre si de 25 dias. Na Praia do Forte, o período internidal médio é de aproximadamente 14.7 dias, e para esta tartaruga específica os outros períodos internidais observados em 87/88 foram de 15, 13 e 13 dias. Para esta mesma tartaruga foram registradas duas desovas na campanha 89/90, com período internidal estimado de 14 dias. Pesquisas referentes a outras populações de tartarugas *Caretta caretta* em locais onde o trabalho de campo teve cobertura praticamente total da praia na temporada de desova indicam que uma *Caretta caretta* raramente deixa de desovar dentro do período internidal esperado (Richardson, 1982; Limpus, 1985).

10.4. Aplicação do modelo aos dados da Praia do Forte: metodologia

10.4.1. Situação dos dados da Praia do Forte em relação à Figura 6.12

A Figura 6.12 fornece, em função de T , uma estimativa do esforço de amostragem necessário para se obter um nível de precisão especificado na estimação de μ . Para se utilizar esta figura no planejamento de trabalho na

praia que tenha por objetivo a obtenção de dados para estimação do número de tartarugas, é necessário uma estimativa prévia (aproximada) de T .

No caso dos dados da Praia do Forte, como o trabalho na praia já foi realizado e o valor $\hat{p}$ já é conhecido a cada temporada (Tabela 10.1), pode-se usar a Figura 6.12 para se situar os dados de cada temporada quanto à precisão na estimação de μ que se pode obter com sua utilização. Para se sobrepor os dados (os valores de $\hat{p}$) da Tabela 10.1 à Figura 6.12, é necessário no entanto alguma estimativa prévia do valor de T para cada temporada. Esta estimativa pode ser obtida com a utilização de valores para μ existentes na literatura, relativos às tartarugas *Caretta caretta*, em estudos onde tenha ocorrido uma cobertura praticamente total da praia.

A distribuição RI7, publicada por Richardson (1982), é uma distribuição obtida de dados reais para uma população de tartarugas *Caretta caretta* nos Estados Unidos (Capítulo 3). A média desta distribuição é 3.32 ninhos por tartaruga. Limpus (1985) obteve, para uma população de *Caretta caretta* na praia de Mon Repos, Austrália, com dados obtidos entre 1977 e 1980, uma média de 3.41 ninhos por tartaruga, bastante próxima à média obtida por Richardson.

A Figura 10.2 mostra os dados da Praia do Forte superpostos à Figura 6.12. Nesta figura, cada um dos oito quadrados (com coordenadas da forma $T \times \hat{p}$) representa uma temporada de desova entre 1987/1988 e 1994/1995, com $\hat{p}$ obtido da Tabela 10.1. Para a obtenção do T aproximado de cada temporada foi utilizada a equação (5.1), dividindo-se o número total de ninhos na temporada (Tabela 10.1) por 3.32, valor tomado como sendo a média do número de ninhos por tartaruga, obtido de Richardson (1982). Observa-se na Figura 10.2 que apenas duas das campanhas, que são as campanhas 87/88 e 90/91, têm um valor

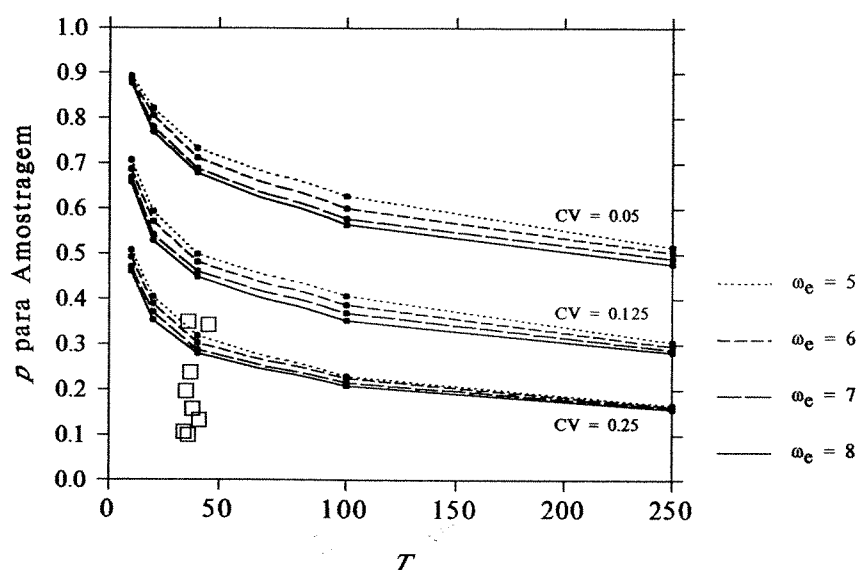


Figura 10.2. Dados das tartarugas *Caretta caretta* da Praia do Forte superpostos à Figura 6.12. Cada quadrado, representando uma temporada de desova (entre 87/88 e 94/95), tem como coordenadas o valor de $\hat{p}$ para a temporada (Tabela 10.1) e um valor para T obtido dividindo-se o total de ninhos da temporada (Tabela 10.1) por 3.32 (valor de μ obtido na literatura, veja o texto).

de $\hat{p}$ suficiente para se fazer não mais que uma estimaco grosseira (segundo o critrio do Captulo 6) do valor de μ .

Uma anlise por ms dos dados destas campanhas 87/88 e 90/91 (Tabela 10.3) mostra que a probabilidade de observao dos ninhos no ,  primeira vista, constante ao longo de cada temporada, como requer o modelo do Captulo 5. Para testar a hiptese de que, a cada campanha, as probabilidades de observao em todos os meses so iguais a $\hat{p}$, foi aplicado um teste de permutao (Efron e Tibshirani, 1993, Captulo 15) (tambm chamado de teste de randomizao (Manly, 1991)). Uma estatstica, definida abaixo,  calculada diretamente dos dados observados, e seu valor  comparado com um valor crtico obtido da distribuio estimada da estatstica, obtida com base na hiptese nula,

Tabela 10.3. Número de ninhos e probabilidade estimada de observação de cada ninho, por mês, campanhas 87/88 e 90/91, Praia do Forte, BA, espécie *Caretta caretta*.

mês	87/88		90/91	
	ninhos	probabilidade	ninhos	probabilidade
setembro	9	0.0000	4	0.0000
outubro	42	0.4286	33	0.3333
novembro	57	0.3509	49	0.5306
dezembro	29	0.3448	24	0.2083
janeiro	10	0.3000	9	0.0000
fevereiro	1	0.0000	1	0.0000

calculada sobre uma amostra de permutações feitas sobre os dados observados, sendo que em cada permutação os ninhos observados e os não-observados são aleatoriamente alocados entre os vários meses e que em cada mês o número total de ninhos no mês é igual ao total real de ninhos no mês. A estatística utilizada (Fleiss, 1981) foi

$$\chi^2 = \sum_{k=\text{setembro}}^{\text{fevereiro}} \frac{N_k(\hat{p}_k - \hat{p})^2}{\hat{p}(1-\hat{p})}, \quad (10.1)$$

onde N_k é o número de ninhos depositados no mês k , $k = \text{setembro}, \dots, \text{fevereiro}$ (Tabela 10.3), $\hat{p}_k$ é a probabilidade estimada de observação a cada mês (Tabela 10.3) e $\hat{p} = \sum_k N_k \hat{p}_k / N$ é a probabilidade estimada de observação para o total da campanha (Tabela 10.1).

O teste, feito com 5000 permutações, não rejeitou a hipótese de igualdade entre os $\hat{p}_k$ para a campanha 87/88, a um nível de significância de 0.05.

Contudo, para a campanha 90/91 houve um resultado altamente significativo

(nível de significância < 0.01), o que indica que os $\hat{p}_k$ não são constantes entre os meses nesta campanha.

10.4.2. Simulação do processo observação-estimação com os $\hat{p}_k$ variando no tempo

Para investigar o impacto que esta variação entre os $\hat{p}_k$ possa ter na estimação de μ , feita através do modelo do Capítulo 5, foram realizadas simulações do processo observação-estimação, nos moldes das simulações da Seção 6.1. Desta vez, no entanto, as simulações tiveram por base populações de tartarugas sobre as quais se impôs uma distribuição temporal na deposição dos ninhos, e o processo de observação foi feito utilizando-se os $\hat{p}_k$ da campanha 90/91, ou seja, a probabilidade de observação dos ninhos variou no tempo. A estimação, contudo, foi feita com o valor $\hat{p}$ da campanha 90/91, conforme requer o modelo.

As simulações foram feitas da seguinte forma:

1. Determinação da distribuição temporal das desovas das tartarugas

Caretta caretta na Praia do Forte: com os dados das campanhas 87/88 a 94/95, foi obtida uma distribuição temporal das desovas na Praia do Forte por períodos de 15 dias aproximadamente, do seguinte modo: o total de desovas no período 1-15 de setembro de todas as campanhas foi dividido pelo total de desovas nestas campanhas, obtendo-se assim a probabilidade estimada de uma desova ser feita no período 1-15 de setembro de qualquer ano. O mesmo foi feito para os outros períodos obtidos pela divisão de cada mês na metade, entre setembro e fevereiro, obtendo-se assim uma distribuição temporal ao longo de 12 períodos, correspondentes a 6 meses. Observe-se que 15 dias é aproximadamente o período

internidal para as tartarugas *Caretta caretta* (Capítulo 2), o que justifica o fato de que, no procedimento descrito abaixo (itens 2a - 2b), cada tartaruga só poderá realizar uma única desova a cada período.

2. De posse da distribuição temporal da Praia do Forte, foram construídas quatro populações de tartarugas tendo as desovas seguindo aproximadamente esta distribuição temporal. Estas populações, com $T = 40$, tiveram por base as populações-modelo com $T = 40$ seguindo as distribuições FL7, RI7, SR7 e ST7, contidas em arquivos de dados (Capítulo 3), e foram construídas do seguinte modo: (a) Para cada tartaruga da população (isto é, para cada tartaruga no arquivo de dados), foi escolhido aleatoriamente um período como sendo o período inicial de desova para esta tartaruga, condicionado a que o total de desovas da tartaruga ocorra dentro dos 12 períodos do item 1. (b) Nos períodos imediatamente seguintes a tartaruga realizou suas demais desovas, uma em cada período, de acordo com seu total de desovas contido no arquivo de dados. Fica assim definida uma configuração temporal para as desovas das 40 tartarugas da população. (c) Foi determinada então a distribuição temporal das desovas para o conjunto das 40 tartarugas, considerando-se os 12 períodos do item 1, e foi calculada a distância entre esta distribuição e a distribuição temporal do item 1, considerando-se cada distribuição como um ponto no espaço real R^{12} dotado da norma euclidiana (veja o Capítulo 3). (d) Os itens 2a - 2c foram repetidos 100 milhões de vezes, e a configuração temporal do item 2b que produziu a menor distância (segundo o item 2c), ou seja, que produziu uma distribuição temporal para o conjunto das tartarugas o mais próximo da distribuição da Praia do Forte, foi escolhida como sendo a configuração a ser utilizada na simulação do processo observação-estimação.

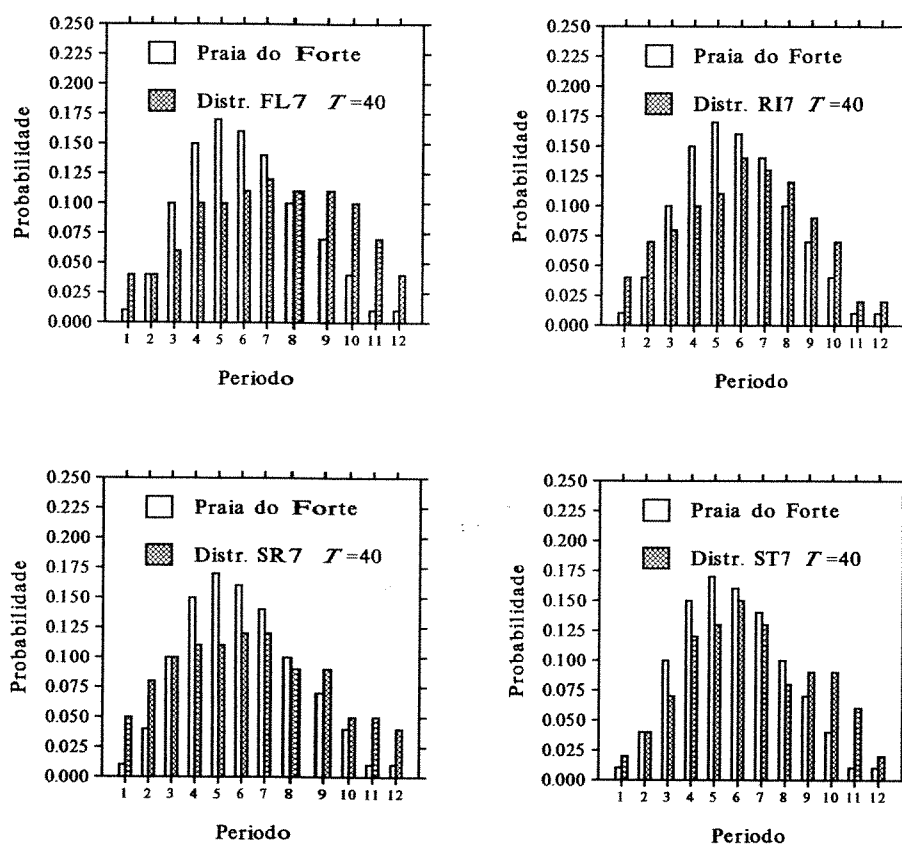


Figura 10.3. Distribuições temporais simuladas das desovas das populações seguindo as distribuições-modelo FL7, RI7, SR7 e ST7 com $T = 40$, comparadas com a distribuição temporal das desovas das tartarugas *Caretta caretta* da Praia do Forte, BA (veja o texto).

Estas construções ocuparam cerca de 24 horas de tempo de processador no sistema IBM SP2. A Figura 10.3 mostra as distribuições temporais obtidas para cada população, comparadas com a distribuição temporal da Praia do Forte.

3. Tendo agora as quatro populações com sua distribuição temporal das desovas, foi feita para cada população uma simulação do processo observação-estimação similar às simulações da Seção 6.1, com 2000 repetições deste processo. No item 3.1 do algoritmo da Seção 6.1, no entanto, Y_j foi computado atribuindo-se a cada ninho da tartaruga j , $j = 1, \dots, 40$, uma probabilidade de

observação correspondente ao mês em que a desova foi realizada, usando-se os $\hat{p}_k$ da campanha 90/91 apresentados na Tabela 10.3 (os dois períodos de aproximadamente 15 dias em cada mês ficam com a mesma probabilidade de observação). Uma vez computados desta forma os Y_j , o algoritmo da Seção 6.1 (que está sendo aplicado para um único valor de p , $p = \hat{p}$) prossegue com a estimação (item 3.4, Seção 6.1), onde se usa o $\hat{p}$ da campanha 90/91, tendo sido utilizado $\omega_e = 7$. Da distribuição dos 2000 $\hat{\mu}$ calculados com este $\hat{p}$ no item 3.4 do algoritmo da Seção 6.1 foram calculados a média e os percentis 2.5 e 97.5.

A Figura 10.4 mostra os resultados destas simulações, comparados com as simulações feitas segundo o método da Seção 6.1, onde as observações dos ninhos ocorreram com o mesmo p atribuído a cada ninho, para diversos valores de p . Vê-se na Figura 10.4 que a observação dos ninhos com os $\hat{p}_k$ variáveis a cada mês, segundo o padrão da campanha 90/91, tende a produzir uma estimativa de μ um pouco maior do que ocorreria se a observação fosse realizada com os $\hat{p}_k$ iguais entre si. O vício relativo (veja o Capítulo 6) para as quatro populações é da ordem de 0.20, bem maior do que o vício relativo para $T = 40$ correspondente a $CV = 0.25$ para a situação em que p é constante na temporada (Figura 6.9), que é da ordem de 0.03. Todo o intervalo de confiança para a estimativa, conforme indicado pelos percentis da distribuição dos 2000 valores de $\hat{\mu}$ das simulações, estará um pouco deslocado para a direita.

10.4.3. Discussão

A análise dos dados da Praia do Forte feita até este ponto mostra a necessidade de se examinar os dados quanto à constância da probabilidade de observação dos ninhos ao longo da temporada, requerida pelo modelo do Capítulo 5. Na realidade, foi realizada, de modo informal, uma análise de

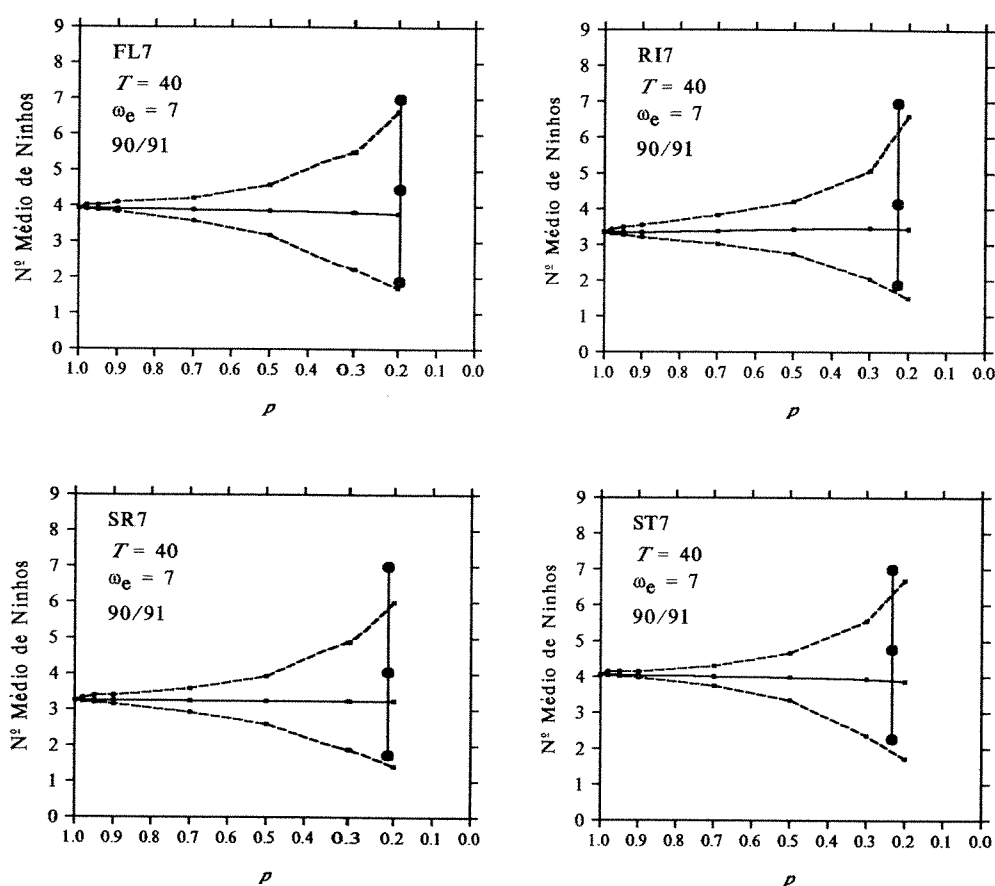


Figura 10.4. Círculos pretos, conectados por linhas verticais: média e percentis 2.5 e 97.5 da distribuição dos 2000 $\hat{\mu}$ calculados a partir das populações com $T = 40$ seguindo as distribuições-modelo FL7, RI7, SR7 e ST7, sobre as quais foi imposta uma distribuição temporal para as desovas. Resultados obtidos por simulação. A probabilidade de observação dos ninhos variou a cada mês, conforme os dados da campanha 90/91. As estimações foram feitas sempre com o $\hat{p}$ da campanha 90/91 e $\omega_e = 7$. Os resultados destas simulações foram superpostos a figuras correspondentes à Figura 6.4 para as distribuições-modelo FL7, RI7, SR7 e ST7.

sensibilidade do estimador do Capítulo 5 em relação à variação de p ao longo da temporada, para o caso específico de uma população de tartarugas marinhas com as desovas seguindo a distribuição temporal das desovas das tartarugas *Caretta caretta* na Praia do Forte. Uma análise mais completa da sensibilidade do estimador em relação à variação de p no tempo teria que lidar com diferentes

distribuições temporais para as desovas e diferentes padrões de variação no tempo para p .

O modo de determinação de uma distribuição temporal para as populações investigadas, através da procura aleatória de uma configuração temporal para as desovas que produzisse uma distribuição temporal o mais próximo possível da distribuição temporal real da Praia do Forte (itens 2a - 2b, Seção 10.4.2) foi uma forma de lidar com a dependência no tempo que existe entre as desovas de uma mesma tartaruga. Como as desovas têm que ocorrer em períodos (aproximados) de 15 dias, não seria possível apenas repartir o total de desovas da temporada entre os 12 períodos de tempo considerados segundo a distribuição temporal real da Praia do Forte, pois neste caso poderiam resultar, para uma mesma tartaruga, períodos internidais de 30, 45, 60 ou mais dias, o que não se admite, em geral, como ocorrendo na natureza para as tartarugas *Caretta caretta* (veja a discussão no final da Seção 10.3).

10.5. Aplicação do modelo aos dados da Praia do Forte: resultados e discussão

Os resultados da seção anterior mostram que, com reservas, é possível aplicar o estimador do Capítulo 5 ao dados da campanha 90/91. Além desta campanha, o estimador, de acordo com a Figura 10.2 e os testes da Seção 10.4, foi aplicado também aos dados da campanha 87/88. As estimações foram feitas com $\omega_e = 7$. Os *bootstraps* para a construção do intervalo de confiança e para o teste do modelo utilizaram 1500 amostras *bootstrap* cada, e o teste do modelo foi feito com $\alpha = 0.05$. Os resultados estão apresentados na Tabela 10.4, que repete alguns dados de tabelas anteriores. Cálculos feitos com $\omega_e = 6$ produziram

Tabela 10.4. Resultados da aplicação do modelo do Capítulo 5 a dados das campanhas 87/88 e 90/91, tartarugas *Caretta caretta*, Praia do Forte, BA.

	campanha	
	87/88	90/91
média de ninhos observada	1.89	1.40
tartarugas observadas	27	30
$\hat{p}$	0.3423	0.3500
teste (Capítulo 9)	0.0279	0.0140
valor critico para o teste, $\alpha = 0.05$	0.1309	0.0595
média de ninhos estimada, $\hat{\mu}$	4.38	2.01
intervalo a 95% para μ	3.02 - 6.05	1.48 - 3.09
tartarugas estimadas, $\hat{T}$	34	60
intervalo a 95% para T	25 - 49	39 - 81

resultados praticamente iguais, em relação à estimação de μ e T , aos da Tabela 10.4.

Pela Tabela 10.4, vê-se que o teste do Capítulo 9 não rejeita o modelo mesmo no caso da campanha 90/91.

O limite inferior do intervalo de confiança para T para a campanha 87/88, 25 tartarugas, é menor do que o número de tartarugas observadas na campanha, 27 tartarugas. Isto sugere que talvez se possa modificar o procedimento referente à construção do intervalo de confiança para T (veja a Seção 8.1), de modo a se utilizar nesta construção a informação disponível sobre o número de tartarugas observadas na temporada.

Observa-se na Tabela 10.4 que embora os dados das duas campanhas sejam, em princípio, provenientes de uma mesma população de tartarugas marinhas (a população de *Caretta caretta* que desova na Praia do Forte), houve uma grande

diferença entre as estimativas $\hat{\mu}$ produzidas a cada campanha, o que é traduzido em intervalos de confiança para μ que mal se tocam. Dado o baixo valor para $\hat{p}$ a cada campanha, é de se esperar uma alta variabilidade na estimação de μ mesmo a partir de dados, a cada campanha, vindos de uma mesma população de tartarugas.

Observa-se também na tabela que os intervalos de confiança para μ e para T não são simétricos em relação às estimativas $\hat{\mu}$ e $\hat{T}$, respectivamente, o que é decorrência do método de construção do intervalo de confiança pelo *bootstrap*, que não impõe condições de simetria ao intervalo construído.

O intervalo de confiança a 95% para o número médio de ninhos na temporada para a campanha 90/91, ao contrário daquele obtido para a campanha 87/88, não cobre os valores encontrados por Richardson (1982) (veja o Capítulo 3) e Limpus (1985) para o número médio de ninhos depositados por uma tartaruga *Caretta caretta* em uma temporada.

CAPÍTULO 11

Conclusões

O modelo apresentado neste trabalho admite extensões, que seriam na realidade outros modelos, aplicáveis a diferentes situações de estimação do número de tartarugas marinhas desovando em uma temporada:

1. Na equação (5.1), N é suposto conhecido. Esta equação é um caso particular da equação (4.1), na qual N em princípio necessita ser também estimado. O modelo pode ser estendido a esta situação em que N deve ser também estimado. Neste caso, para o cálculo de um intervalo de confiança por um *bootstrap*, seria necessário dispor de uma estimativa para a distribuição amostral de $\hat{N}$.
2. A distribuição $\hat{F}$ definida por (5.12), calculada a partir de dados de marcação obtidos em uma praia determinada, pode ser utilizada como estimativa para a distribuição F do número de ninhos por tartaruga para toda uma região vizinha à praia de onde se obteve $\hat{F}$. Isto permitiria utilizar os dados de marcação obtidos em uma única praia, ou mesmo em um trecho de praia, como base para a estimação do número de tartarugas marinhas desovando em todo um conjunto de praias, dado que se conhecesse o valor de N para este conjunto de praias.
3. Conjugando-se os itens 1 e 2 acima, seria obtido um modelo que possibilitaria a estimação do número de fêmeas desovando na temporada em toda

uma região, com a obtenção de dados por marcação ocorrendo apenas em uma praia ou em um determinado trecho de praia e com o número total de ninhos na região na temporada sendo estimado por um processo de amostragem.

Situações de pesquisa em que N é estimado por meios como inquéritos aéreos e/ou por trabalhos em terra com cobertura incompleta da praia são encontradas na literatura (Carr *et al.*, 1978; Shoop *et al.*, 1985; Mortimer, 1988; Hopkins-Murphy e Murphy, 1994). O trabalho de Carr *et al.* (1978) na praia de Tortuguero, Costa Rica, exemplifica a situação em que a atividade de marcação é realizada apenas em um trecho de praia.

Da mesma forma do que ocorre no uso de outros métodos de estimação de tamanho de populações animais (Lebreton *et al.*, 1992), a utilização do modelo proposto no presente trabalho para as tartarugas marinhas requer o uso de programa de computador, devido à complexidade e à quantidade dos cálculos a serem realizados.

A análise dos dados do Projeto TAMAR (Capítulo 10) mostra a utilidade que pode ter a Figura 6.12 no planejamento de ações de campo que tenham entre os seus objetivos a estimação da população de tartarugas marinhas desovando a cada temporada. Esta figura indica o esforço de amostragem necessário para a obtenção de um nível especificado de precisão na estimação da população de fêmeas desovando em uma temporada, na situação em que todos os ninhos da temporada sejam contados. A análise do Capítulo 10 ressalta também a importância de se realizar um trabalho de campo que produza um valor para p constante (aproximadamente) ao longo da temporada, conforme requerido para a aplicação do modelo de estimação do Capítulo 5.

APÊNDICE A

Distribuições-modelo para o número de ninhos por tartaruga para $\omega_d = 5, 6$ e 8

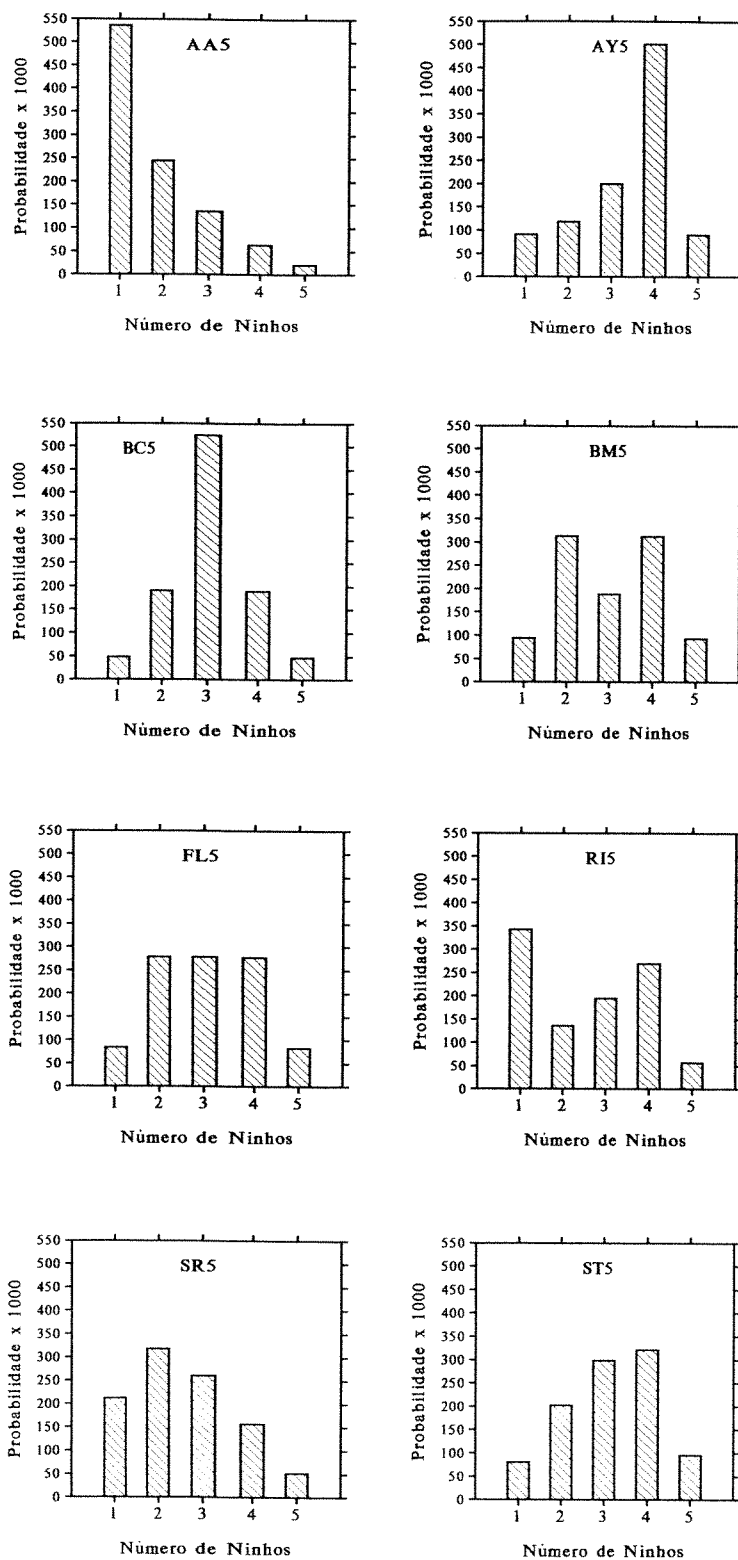


Figura A.1. Distribuições-modelo para o número de ninhos por tartaruga com $\omega_d = 5$: AA5, AY5, BC5, BM5, FL5, RI5, SR5 e ST5.

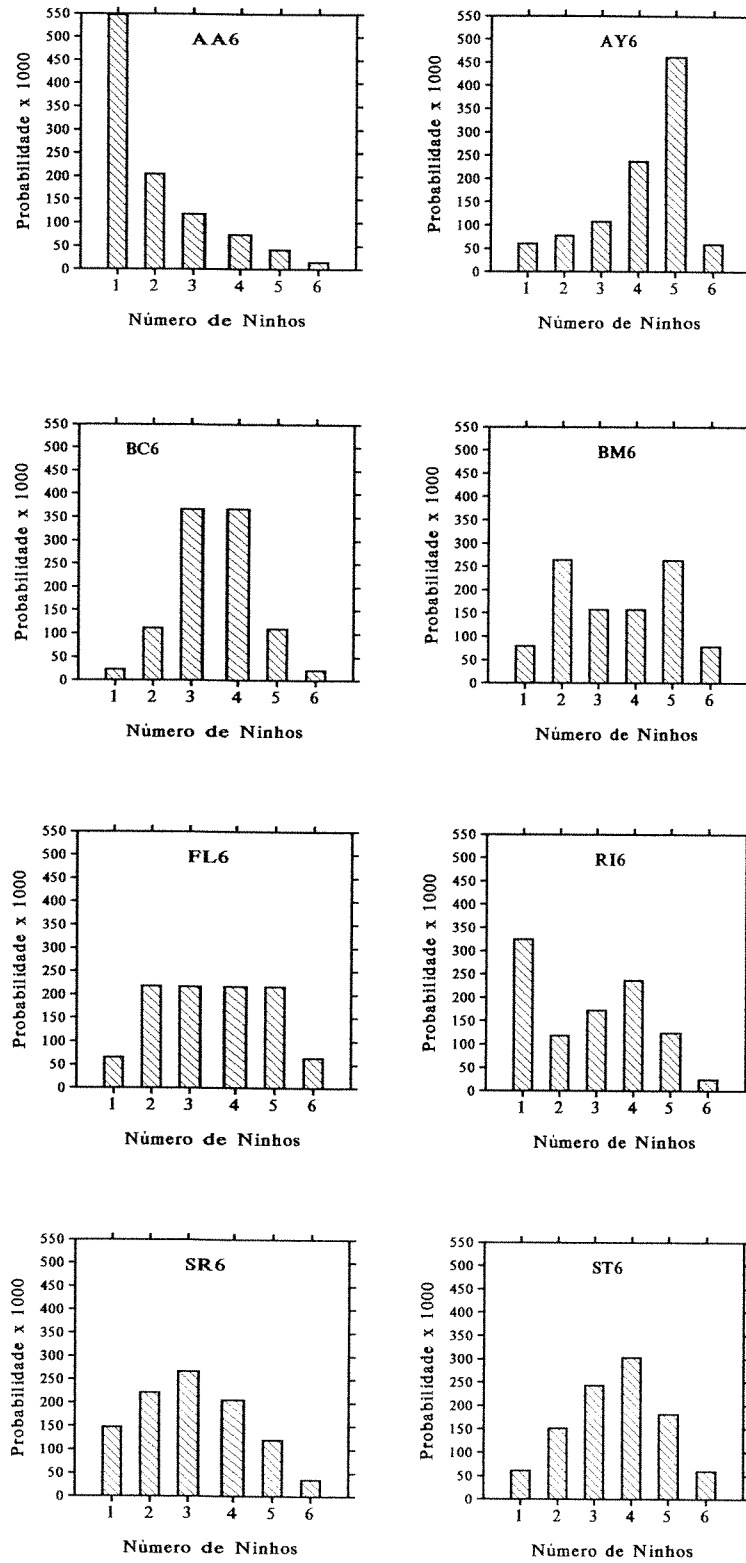


Figura A.2. Distribuições-modelo para o número de ninhos por tartaruga com $\omega_d = 6$: AA6, AY6, BC6, BM6, FL6, RI6, SR6 e ST6.

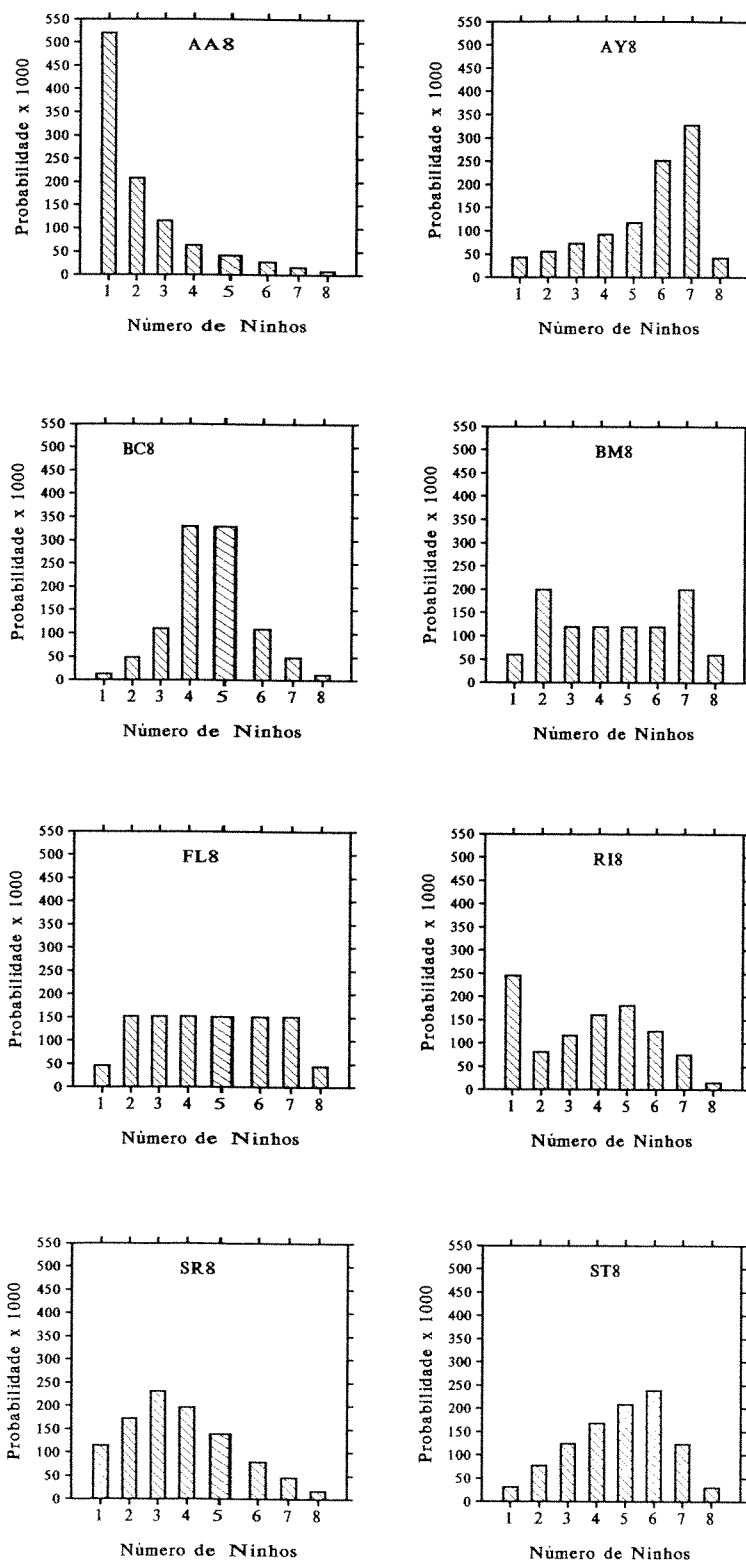


Figura A.3 Distribuições-modelo para o número de ninhos por tartaruga com $\omega_d = 8$: AA8, AY8, BC8, BM8, FL8, RI8, SR8 e ST8.

APÊNDICE B

Resultados adicionais da avaliação do método de construção do intervalo de confiança

Neste apêndice são apresentados resultados análogos aos da Figura 8.5 para $\omega_e = 5, 6$ e 8 , relativos ao método para a construção do intervalo de confiança por meio de um *bootstrap*.

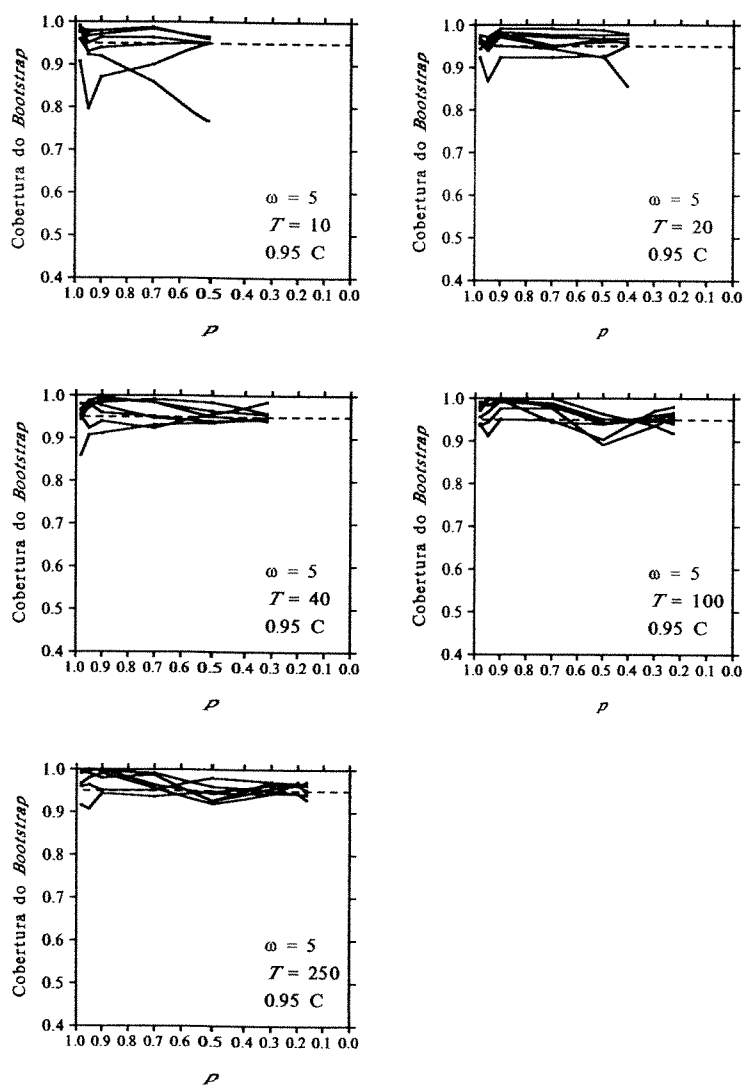


Figura B.1. Cobertura do intervalo de confiança produzido pelo *bootstrap* em função de p , resultados obtidos por simulação para $\omega_e = 5$, com calibração. Cada linha cheia representa uma população-modelo (veja o texto). $T = 10, 20, 40, 100$ e 250 . A linha horizontal tracejada indica a cobertura nominal, igual a 0.95. Os resultados são apresentados apenas na região útil (aproximadamente) para cada T . Observar a escala reversa no eixo horizontal de cada gráfico.

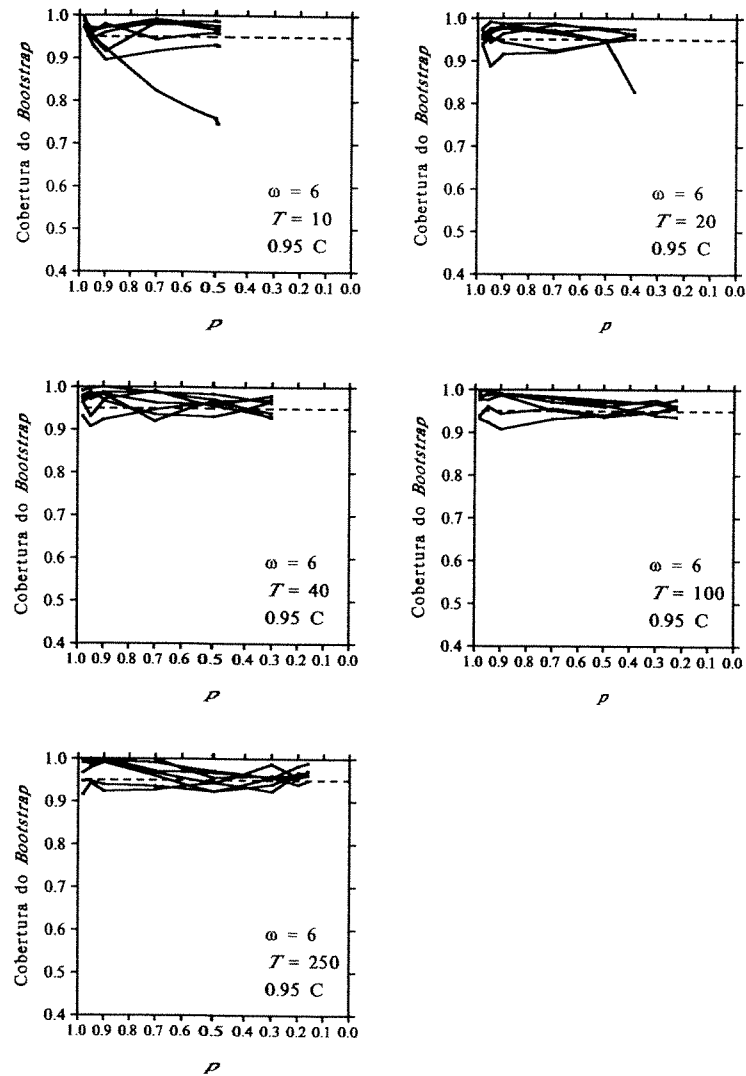


Figura B.2. Cobertura do intervalo de confiança produzido pelo *bootstrap* em função de p , resultados obtidos por simulação para $\omega_e = 6$, com calibração. Cada linha cheia representa uma população-modelo (veja o texto). $T = 10, 20, 40, 100$ e 250 . A linha horizontal tracejada indica a cobertura nominal, igual a 0.95 . Os resultados são apresentados apenas na região útil (aproximadamente) para cada T . Observar a escala reversa no eixo horizontal de cada gráfico.

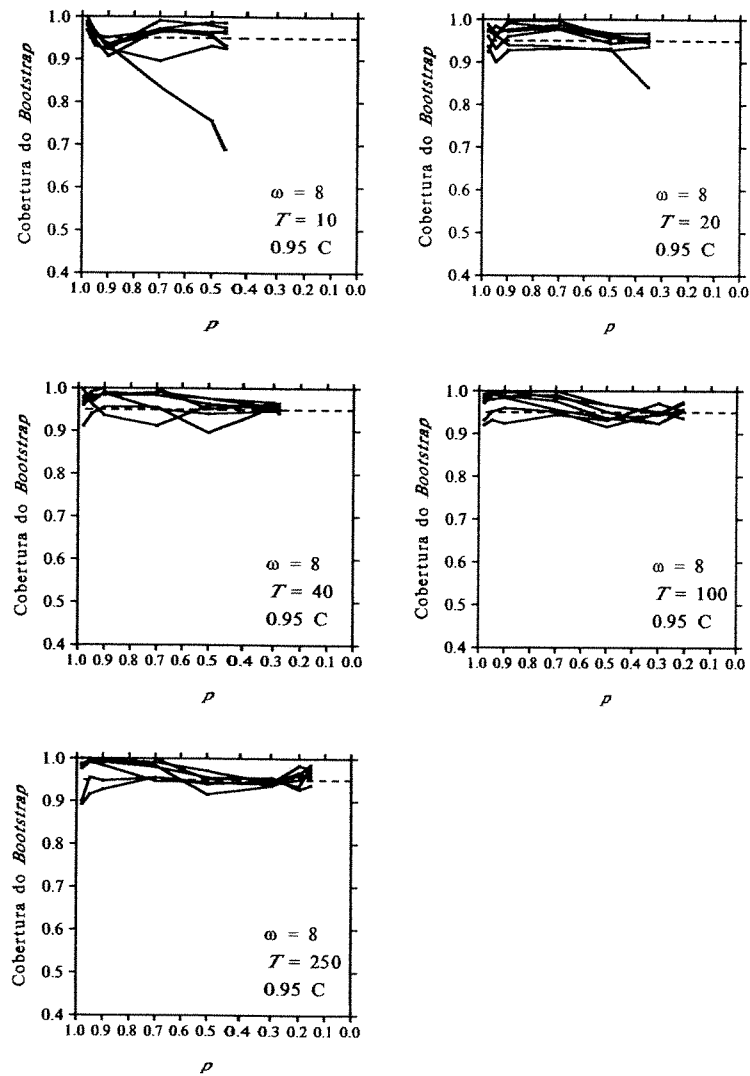


Figura B.3. Cobertura do intervalo de confiança produzido pelo *bootstrap* em função de p , resultados obtidos por simulação para $\omega_e = 8$, com calibração. Cada linha cheia representa uma população-modelo (veja o texto). $T = 10, 20, 40, 100$ e 250 . A linha horizontal tracejada indica a cobertura nominal, igual a 0.95. Os resultados são apresentados apenas na região útil (aproximadamente) para cada T . Observar a escala reversa no eixo horizontal de cada gráfico.

APÊNDICE C

Resultados adicionais da avaliação do teste do modelo

Neste apêndice são apresentados resultados análogos aos da Figura 9.3 para $\omega_e = 5, 6$ e 8 , relativos ao teste do modelo.

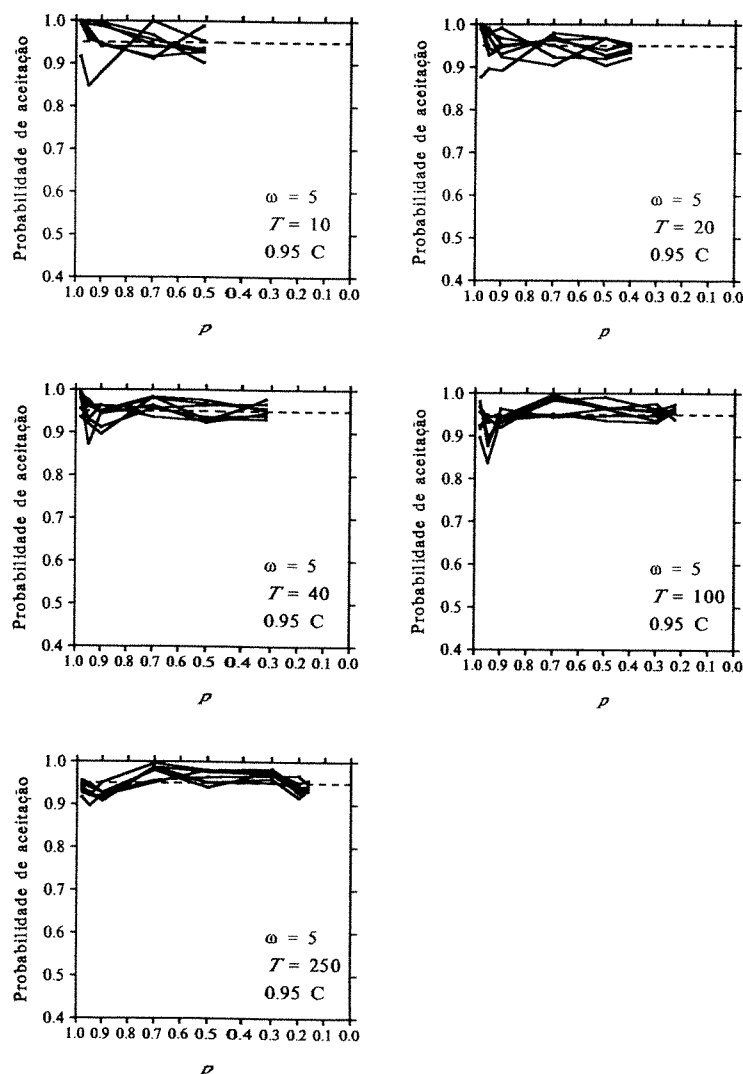


Figura C.1. Probabilidade de aceitação do teste, em função de p , resultados obtidos por simulação para $\omega_e = 5$, com calibração. Cada linha cheia representa uma população-modelo (veja o texto). $T = 10, 20, 40, 100$ e 250 . A linha horizontal tracejada indica $1 - \alpha = 0.95$, onde $\alpha = 0.05$ é o erro de tipo I desejado para o teste. Os resultados são apresentados apenas na região útil (aproximadamente) para cada T . Observar a escala reversa no eixo horizontal de cada gráfico.

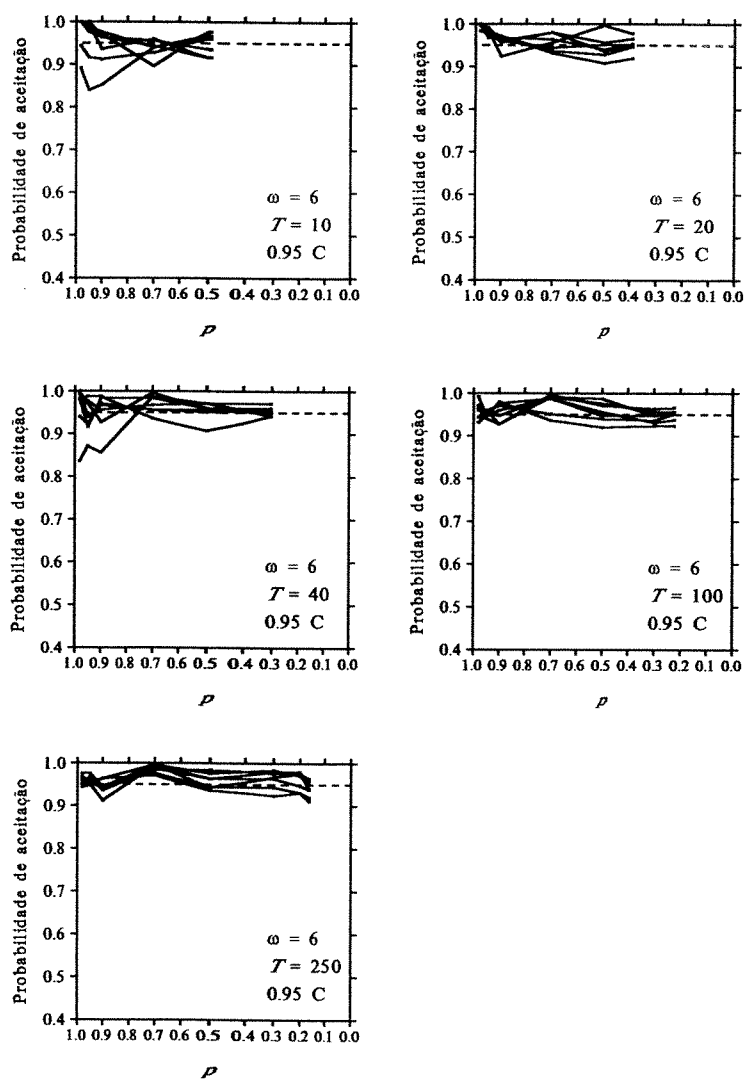


Figura C.2. Probabilidade de aceitação do teste, em função de p , resultados obtidos por simulação para $\omega_e = 6$, com calibração. Cada linha cheia representa uma população-modelo (veja o texto). $T = 10, 20, 40, 100$ e 250 . A linha horizontal tracejada indica $1 - \alpha = 0.95$, onde $\alpha = 0.05$ é o erro de tipo I desejado para o teste. Os resultados são apresentados apenas na região útil (aproximadamente) para cada T . Observar a escala reversa no eixo horizontal de cada gráfico.

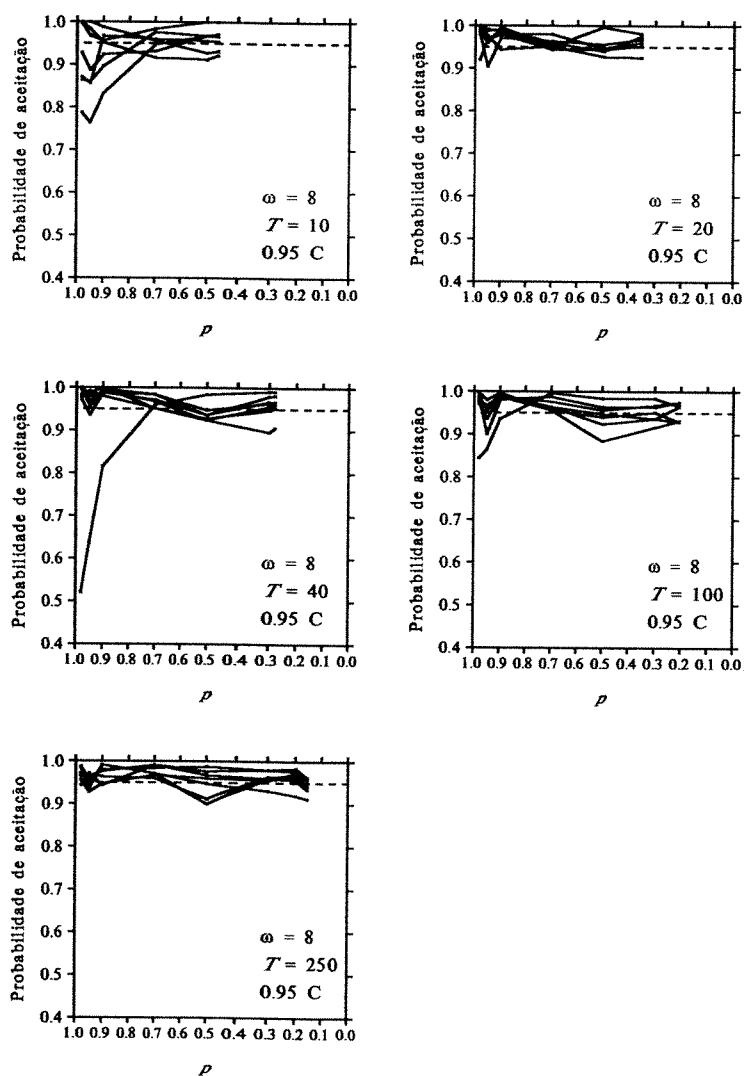


Figura C.3. Probabilidade de aceitação do teste, em função de p , resultados obtidos por simulação para $\omega_e = 8$, com calibração. Cada linha cheia representa uma população-modelo (veja o texto). $T = 10, 20, 40, 100$ e 250 . A linha horizontal tracejada indica $1 - \alpha = 0.95$, onde $\alpha = 0.05$ é o erro de tipo I desejado para o teste. Os resultados são apresentados apenas na região útil (aproximadamente) para cada T . Observar a escala reversa no eixo horizontal de cada gráfico.

Referências bibliográficas

- AVRIEL, M., W.E. DIEWERT, S. SCHAIBLE e W.T. ZIEMBA. 1981.
Introduction to concave and generalized concave functions. Páginas 21-50 *in* S. Schaible e W.T. Ziemba (editores). Generalized Concavity in Optimization and Economics. Academic Press, New York, USA.
- BAZARAA, M.S., H.D. SHERALI e C.M. SHETTY. 1993. Nonlinear Programming: Theory and Algorithms. Segunda edição. Wiley, New York, USA.
- BEGON, M., J.L. HARPER e C.R. TOWNSEND. 1990. Ecology: Individuals, Populations and Communities. Segunda edição. Blackwell, Boston, USA.
- BERAN, R. 1984. Minimum distance procedures. Páginas 741-754 *in* P. R. Krishnaiah e P.K. Sen (editores). Handbook of Statistics, volume 4. Elsevier, Amsterdam, Holanda.
- BERAN, R. e P.W. MILLAR. 1987. Stochastic estimation and testing. The Annals of Statistics **15**: 1131-1154.
- BOLTEN, A.B., K.A. BJORNDAL e H.R. MARTINS. 1994. Biology of pelagic-stage loggerheads in the Atlantic. Páginas 19-20 *in* K.A. Bjorndal, A.B. Bolten, D.A. Johnson e P.J. Eliazar (organizadores). Proceedings of the Fourteenth Annual Symposium on Sea Turtle Biology and Conservation. NOAA Technical Memorandum NMFS-SEFSC-351, USA.
- BORLAND INTERNATIONAL, INC. 1992. Turbo Pascal Version 7.0. Language Guide. Borland, USA.

- BOULON, R.H., JR., D.L. McDONALD e P.H. DUTTON. 1994. Leatherback turtle (*Dermochelys coriacea*) nesting biology, Sandy Point, St. Croix, U.S. Virgin Islands: 1981-1993. Páginas 190-193 in K.A. Bjorndal, A.B. Bolten, D.A. Johnson e P.J. Eliazar (organizadores). Proceedings of the Fourteenth Annual Symposium on Sea Turtle Biology and Conservation. NOAA Technical Memorandum NMFS-SEFSC-351, USA.
- BOWEN, B.W., A.B. MEYLAN, J.P. ROSS, C.J. LIMPUS, G.H. BALAZS e J.C. AVISE. 1992. Global population structure and natural history of the green turtle (*Chelonia mydas*) in terms of matriarchal phylogeny. *Evolution* **46**: 865-881.
- BOWEN, B.W., N. KAMEZAKI, C.J. LIMPUS, G.R. HUGHES, A.B. MEYLAN, e J.C. AVISE. 1994. Global phylogeography of the loggerhead turtle (*Caretta caretta*) as indicated by mitochondrial DNA haplotypes. *Evolution* **48**: 1820-1828.
- BUSKIRK, J.V. e L.B. CROWDER. 1994. Life-history variation in marine turtles. *Copeia* **1994**: 66-81.
- BUSTARD, H.R. 1979. Population dynamics of sea turtles. Páginas 523-540 in M. Harless e H. Morlock (editores). *Turtles: Perspectives and Research*. Wiley, New York, USA.
- CARR, A. 1975. The Ascension Island green turtle colony. *Copeia* **1975**: 547-555.
- CARR, A. 1980. Some problems of sea turtle ecology. *American Zoologist* **20**: 489-498.
- CARR, A. 1987. New perspectives on the pelagic stage of sea turtle development. *Conservation Biology* **1**: 103-121.

- CARR, A. e H. HIRTH. 1962. The ecology and migrations of sea turtles, 5. Comparative features of isolated green turtle colonies. American Museum Novitates **2091**: 1-42.
- CARR, A., M.H. CARR e A.B. MEYLAN. 1978. The ecology and migrations of sea turtles, 7. The West Caribbean green turtle colony. Bulletin of the American Museum of Natural History **162**: 1-46.
- CASELLA, G. e R.L. BERGER. 1990. Statistical Inference. Duxbury Press / Wadsworth, Belmont, California, USA.
- COMPUTER ASSOCIATES INTERNATIONAL, INC. 1992. CA-Clipper Reference Guide, Version 5.2, volumes 1 e 2. Computer Associates, USA.
- CROUSE, D.T., L.B. CROWDER e H. CASWELL. 1987. A stage-based population model for loggerhead sea turtles and implications for conservation. Ecology **68**: 1412-1423.
- DENNIS, B., P.L. MUNHOLLAND e J.M. SCOTT. 1991. Estimation of growth and extinction parameters for endangered species. Ecological Monographs **61**: 115-143.
- DIEWERT, W.E., M. AVRIEL e I. ZANG. 1981. Nine kinds of quasiconcavity and concavity. Journal of Economic Theory **25**: 397-420.
- DODD, C.K., Jr. 1988. Synopsis of the Biological Data on the Loggerhead Sea Turtle *Caretta caretta* (Linnaeus 1758). U.S. Fish and Wildlife Service, Biological Report **88**(14), Washington, D.C., USA.
- EFRON, B. 1982. The Jackknife, the Bootstrap and Other Resampling Plans. Volume 38 of CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, USA.

- EFRON, B. e R.J. TIBSHIRANI. 1993. An Introduction to the Bootstrap. Chapman & Hall, New York, USA.
- EHRHART, L.M. 1982. A review of sea turtle reproduction. Páginas 29-38 in K. Bjorndal (editora). Biology and Conservation of Sea Turtles. Smithsonian Institution Press, Washington, DC, USA.
- EMLLEN, J.M. 1984. Population Biology. Macmillan, New York, USA.
- FLEISS, J.L. 1981. Statistical Methods for Rates and Proportions. Segunda edição. Wiley, New York, USA.
- FOSTER, K. 1994. A growth curve for wild Florida *Caretta caretta*. Páginas 221-224 in K.A. Bjorndal, A.B. Bolten, D.A. Johnson e P.J. Eliazar (organizadores). Proceedings of the Fourteenth Annual Symposium on Sea Turtle Biology and Conservation. NOAA Technical Memorandum NMFS-SEFSC-351, USA.
- FRAZER, N.B. 1983. Survivorship of adult female loggerhead sea turtles, *Caretta caretta*, nesting on Little Cumberland Island, Georgia, USA. Herpetologica **39**: 436-447.
- FRAZER, N.B. 1986. Survival from egg to adulthood in a declining population of loggerhead turtles, *Caretta caretta*. Herpetologica **42**: 47-55.
- FRAZER, N.B. e R.C. LADNER. 1986. A growth curve for green turtles, *Chelonia mydas*, in the U.S. Virgin Islands, 1913-14. Copeia **1986**: 798-802.
- FRAZER, N.B. e J.I. RICHARDSON. 1985. Annual variation in clutch size and frequency for loggerhead turtles, *Caretta caretta*, nesting at Little Cumberland Island, Georgia, USA. Herpetologica **41**: 246-251.
- FRAZER, N.B., C.J. LIMPUS e J.L. GREENE. 1994. Growth and estimated age at maturity of Queensland loggerheads. Páginas 42-46 in K.A. Bjorndal, A.B.

- Bolten, D.A. Johnson e P.J. Eliazar (organizadores). Proceedings of the Fourteenth Annual Symposium on Sea Turtle Biology and Conservation. NOAA Technical Memorandum NMFS-SEFSC-351, USA.
- GREEN, D. 1994. Galapagos sea turtles: an overview. Páginas 65-68 in B.A. Schroeder e B.E. Witherington (organizadores). 1994. Proceedings of the Thirteenth Annual Symposium on Sea Turtle Biology and Conservation. NOAA Technical Memorandum NMFS-SEFSC-341, USA.
- GYURIS, E. e C.J. LIMPUS. 1988. The loggerhead turtle, *Caretta caretta*, in Queensland: population breeding structure. Wildlife Research **15**: 197-209.
- HILLESTAD, H.O., J.I. RICHARDSON, C. McVEA, Jr. e J.M. WATSON, Jr. 1982. Worldwide incidental capture of sea turtles. Páginas 489-495 in K. Bjorndal (editora). Biology and Conservation of Sea Turtles. Smithsonian Institution Press, Washington, DC, USA.
- HIRTH, H.F. 1980. Some aspects of the nesting behavior and reproductive biology of sea turtles. American Zoologist **20**: 507-523.
- HOPKINS-MURPHY, S.R. e T.M. MURPHY. 1994. Status of the loggerhead nesting population in South Carolina: a five year up-date. Páginas 62-64 in K.A. Bjorndal, A.B. Bolten, D.A. Johnson e P.J. Eliazar (organizadores). Proceedings of the Fourteenth Annual Symposium on Sea Turtle Biology and Conservation. NOAA Technical Memorandum NMFS-SEFSC-351, USA.
- HUGHES, G.R. 1974a. The sea turtles of South-East Africa. I. Status, morphology and distributions. Investigational Report No. 35. Oceanographic Research Institute, Durban, South Africa.
- HUGHES, G.R. 1974b. The sea turtles of South-East Africa. II. The biology of the Tongaland loggerhead turtle *Caretta caretta* L. with comments on the

leatherback turtle *Dermochelys coriacea* L. and the green turtle *Chelonia mydas* L. in the study region. Investigational Report No. 36. Oceanographic Research Institute, Durban, South Africa.

HUGHES, G.R. 1982. Nesting cycles in sea turtles - typical or atypical? Páginas 81-89 in K. Bjorndal (editora). Biology and Conservation of Sea Turtles. Smithsonian Institution Press, Washington, DC, USA.

INTERNATIONAL BUSINESS MACHINES CORP. 1992. AIX XL Fortran Compiler / 6000. Language Reference, Version 2.3. IBM, USA.

IUCN / SSC. 1995. A Global Strategy for the Conservation of Marine Turtles. IUCN/SSC, USA.

JOHNSON, D.H. 1994. Population Analysis. Páginas 419-444 in T.A. Bookhout (editor). Research and Management Techniques for Wildlife and Habitats. Quinta edição. The Wildlife Society, Bethesda, Maryland, USA.

KREBS, C.J. 1989. Ecological Methodology. Harper & Row, New York, USA.

LEBRETON, J.D., K.P. BURNHAM, J. CLOBERT e D.R. ANDERSON. 1992. Modeling survival and testing biological hypothesis using marked animals: a unified approach with case studies. Ecological Monographs **62**: 67-118.

L'ECUYER, P. 1988. Efficient and portable combined random number generators. Communications of the Association of Computing Machinery **31**: 742-749 e 774.

LE GALL, J.Y., P. BOSC, D. CHATEAU e M. TAQUET. 1986. Estimation du nombre de tortues vertes femelles adultes *Chelonia mydas* par saison de ponte à Tromelin et Europa (Océan Indien) (1973-1985). Océanographie Tropicale **21**: 3-22.

- LENARZ, M.S., N.B. FRAZER, M.S. RALSTON e R. B. MAST. 1981. Seven nests recorded for loggerhead turtle (*Caretta caretta*) in one season. *Herpetological Review* **12**: 9.
- LIMPUS, C.J. 1985. A Study of the Loggerhead Sea Turtle, *Caretta caretta*, in Eastern Australia. Dissertação de Doutorado. University of Queensland, St. Lucia, Australia.
- LIMPUS, C.J. 1992a. The hawksbill turtle, *Eretmochelys imbricata*, in Queensland: population structure within a southern Great Barrier Reef feeding ground. *Wildlife Research* **19**: 489-506.
- LIMPUS, C.J. 1992b. Estimation of tag loss in marine turtle research. *Wildlife Research* **19**: 457-469.
- LIMPUS, C.J. 1994. The loggerhead turtle, *Caretta caretta*, in Queensland: feeding ground selection following her first nesting season. Páginas 78-81 in K.A. Bjorndal, A.B. Bolten, D.A. Johnson e P.J. Eliazar (organizadores). *Proceedings of the Fourteenth Annual Symposium on Sea Turtle Biology and Conservation*. NOAA Technical Memorandum NMFS-SEFSC-351, USA.
- LIMPUS, C.J., P.J. COUPER e K.L.D. COUPER. 1993. Crab Island revisited: reassessment of the world's largest flatback turtle rookery after twelve years. *Memoirs of the Queensland Museum* **33**: 277-289.
- LIMPUS, C.J., P.J. COUPER e M. A. READ. 1994. The green turtle, *Chelonia mydas*, in Queensland: population structure in a warm temperate feeding area. *Memoirs of the Queensland Museum* **35**: 139-154.
- LIMPUS, C.J., J.D. MILLER, C.J. PARMENTER, D. REIMER, N. McLACHLAN e R. WEBB. 1992. Migration of green (*Chelonia mydas*) and

- loggerhead (*Caretta caretta*) turtles to and from eastern Australian rookeries. *Wildlife Research* **19**: 347-358.
- MANLY, B.F.J. 1991. *Randomization and Monte Carlo Methods in Biology*. Chapman & Hall, London, UK.
- MARQUEZ M., R. 1994. Synopsis of Biological Data on the Kemp's Ridley Turtle, *Lepidochelys kempii* (Garman, 1880). NOAA Technical Memorandum NMFS-SEFSC-343, USA.
- MARQUEZ M., R., A. VILLANUEVA O. e M.S. PEREZ. 1982. The population of Kemp's ridley sea turtle in the Gulf of Mexico - *Lepidochelys kempii*. Páginas 159-164 in K. Bjorndal (editora). *Biology and Conservation of Sea Turtles*. Smithsonian Institution Press, Washington, DC, USA.
- MAY, R.M. (editor). 1981. *Theoretical Ecology*. Segunda edição. Sinauer, Sunderland, Massachusetts, USA.
- McNEELY, J.A., K.R. MILLER, W.V. REID, R.A. MITTERMEIER e T.B. WERNER. 1990. *Conserving the World's Biological Diversity*. IUCN, Gland, Switzerland; WRI, CI, WWF-US e The World Bank, Washington, DC, USA.
- MEYLAN, A. 1982. Estimation of population size in sea turtles. Páginas 135-138 in K. Bjorndal (editora). *Biology and Conservation of Sea Turtles*. Smithsonian Institution Press, Washington, DC, USA.
- MORTIMER, J.A. 1988. Green turtle nesting at Aldabra Atoll - population estimates and trends. *Bulletin of the Biological Society of Washington* **8**: 116-128.
- MORTIMER, J.A. e A. CARR. 1987. Reproduction and migrations of the Ascension Island green turtle (*Chelonia mydas*). *Copeia* **1987**: 103-113.

- MROSOVSKY, N. 1983. *Conserving Sea Turtles*. The British Herpetological Society, London, UK.
- NUMERICAL ALGORITHMS GROUP, INC. 1993. *Fortran Library Manual*, Mark 16. NAG, USA.
- OWEN, R.D., S.A. JOHNSON, W.E. REDFOOT e L.M. EHRHART. 1994. Marine turtle nest production and reproductive success at Archie Carr NWR: 1982-1993. Páginas 109-111 in K.A. Bjorndal, A.B. Bolten, D.A. Johnson e P.J. Eliazar (organizadores). *Proceedings of the Fourteenth Annual Symposium on Sea Turtle Biology and Conservation*. NOAA Technical Memorandum NMFS-SEFSC-351, USA.
- PARR, W.C. 1985. Minimum distance estimation. Páginas 529-532 in Kotz, S. e N.L. Johnson (editores). *Encyclopedia of Statistical Sciences*, volume 5. Wiley, New York, USA.
- POLLOCK, K.H., J.D. NICHOLS, C. BROWNIE e J.E. HINES. 1990. Statistical inference for capture-recapture experiments. *Wildlife Monographs* 107: 1-97.
- PRITCHARD, P.C.H. 1979. *Encyclopaedia of Turtles*. T.F.H. Publications, Neptune, New Jersey, USA.
- PRITCHARD, P.C.H. 1980. The conservation of sea turtles: practices and problems. *American Zoologist* 20: 609-617.
- PRITCHARD, P.C.H. e P. TREBBAU. 1984. *The Turtles of Venezuela*. Contributions to Herpetology no. 2. Society for the Study of Amphibians and Reptiles, U.S.A.
- PRITCHARD, P., P. BACON, F. BERRY, A. CARR, J. FLETMEYER, R. GALLAGHER, S. HOPKINS, R. LANKFORD, R. MARQUEZ M., L. OGREN, W. PRINGLE, Jr., H. REICHART e R. WITHAM. 1983. *Manual*

- of Sea Turtle Research and Conservation Techniques. Second edition. K.A. Bjorndal e G.H. Campbell (editores). Center for Environmental Education, Washington, D.C., USA.
- RAFTERY, A.E. e T. SCHWEDER. 1993. Inference about the ratio of two parameters, with application to whale censusing. *The American Statistician* **47**: 259-264.
- REICHART, H.A. 1993. *Synopsis of Biological Data on the Olive Ridley Sea Turtle *Lepidochelys olivacea* (Eschscholtz, 1829) in the Western Atlantic*. NOAA Technical Memorandum NMFS-SEFSC-336, USA.
- RICHARDSON, J.I. 1982. *A Population Model for Adult Female Loggerhead Sea Turtles (*Caretta caretta*) Nesting in Georgia*. Dissertação de doutorado. University of Georgia, Athens, Georgia, USA.
- RICHARDSON, J.I., T.H. RICHARDSON e M.W. DIX. 1978. Population estimates for nesting female loggerhead sea turtles (*Caretta caretta*) in the St. Andrew Sound area of southeastern Georgia. *Florida Marine Research Publications* **33**: 34-38.
- SCHROEDER, B.A. 1994. Florida index nesting beach surveys: are we on the right track? Páginas 132-133 in K.A. Bjorndal, A.B. Bolten, D.A. Johnson e P.J. Eliazar (organizadores). *Proceedings of the Fourteenth Annual Symposium on Sea Turtle Biology and Conservation*. NOAA Technical Memorandum NMFS-SEFSC-351, USA.
- SCHULZ, J.P. 1975. Sea turtles nesting in Surinam. *Zoologische Verhandelingen (Leiden)* **143**: 1-143.
- SEABROOK, W. 1991. How many green turtles nest on Aldabra? *Oryx* **25**: 96-98.

- SHOOP, C.R., C. A. RUCKDESCHEL e N.B. THOMPSON. 1985. Sea turtles in the southeast United States: nesting activity as derived from aerial and ground surveys, 1982. *Herpetologica* **41**: 252-259.
- WITHAM, R. 1980. The "lost year" question in young sea turtles. *American Zoologist* **20**: 525-530.
- WITZELL, W.N. 1983. *Synopsis of Biological Data on the Hawksbill Turtle *Eretmochelys imbricata* (Linnaeus, 1766)*. FAO Fisheries Synopsis 137. Food and Agriculture Organization, Rome, Italy.