**Inferência bayesiana
do número de espécies
de uma população**

Olga Satomi Yoshida

TESE APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA OBTENÇÃO DO GRAU DE DOUTORA
EM
ESTATÍSTICA

Área de Concentração: **Estatística**
Orientador: **Prof. Dr. José Galvão Leite**

Trabalho parcialmente financiado pela CAPES

–São Paulo, dezembro de 1996–

Inferência bayesiana do número de espécies de urna população

Este exemplar corresponde a redação final da tese devidamente corrigida, defendida por Olga Satomi Yoshida e aprovada pela comissão julgadora.

São Paulo, 19 de dezembro de 1996

Banca examinadora:

Prof. Dr. José Galvão Leite (orientador)- IME-USP

Prof. Dr Heleno Bolfarine - IME-USP

Prof. Dr Carlos Daniel Mimoso Paulino- Univ. Tec. de Lisboa

Prof. Dr Jorge Alberto Achcar - ICMSC-USP

Prof. Dr Mauro Sergio de Freitas Marques - IMECC-UNICAMP

À memória do 'ditchan'
Jisaku Yoshida
Japão/1898-Brasil/1945

Agradecimentos

Tive a honra de trabalhar com o professor José Galvao Leite. Além do seu positivismo e alto astral, contei com a sua incansável e ativa participação desde os primeiros seminários até a fase final deste trabalho, mesmo quando estive ausente deste instituto. A sua leitura meticulosa e crítica com inúmeras sugestões melhoraram muito a redação desta tese.

Ao Professor Heleno Bolfarine que no meio de suas incontáveis atividades sempre achou um tempo para participar das nossas discussões dando valiosas sugestões e para auxiliar na atualização das nossas referências bibliográficas.

Durante minha pós-graduação aqui no IME algumas pessoas se tornaram muito especiais para mim. São elas: Pilar Loreto Iglesias Zuazola, Reinaldo Aureliano Valle e Carlos Daniel Mimoso Paulino, por coincidência todos Bayesianos. Eles foram e continuam sendo referências importantes para mim.

A Beti Kira, Antonio, Julia, Geraldo, Lourdes, Helena e Cecilia pela ajuda que me deram com correções de editoração e auxílio técnico.

Ao Fábio e André. Na verdade, qualquer coisa sem os dois não teria a menor graça.

A Mãe e ao Pá.

Sumário

A estimação de k , número desconhecido de categorias ou espécies de uma população, baseada nas frequências das frequências de uma amostra desta população é conhecida por problema das espécies. Há muitas versões do problema das espécies. O número de espécies, k , pode ser, por exemplo, o tamanho do vocabulário de Shakespeare, o número de fontes de erros de um novo sistema de programas para computador, o número de componentes de uma superposição de processos ou o tamanho de uma população. Consideramos formulações Bayesianas dos modelos estatísticos mais frequentes na literatura para resolver este problema. Em todas as formulações, a distribuição a posteriori de k depende do número de espécies distintas observadas na amostra, da distribuição a priori e de uma quantidade determinada pelo planejamento amostral, como por exemplo, tamanho da amostra, tempo de observação da amostra ou número de capturas. Mostramos que a distribuição a posteriori de k é estocasticamente monótona em cada um destes parâmetros e, com isto, descrevemos como os dados e a priori influenciam na estimação de k . Mostramos também que os estimadores de Bayes para k assim como outras quantidades de interesse são consistentes.

Abstract

The estimation of k , the unknown number of categories or species in a population, based on the frequencies of frequencies of a sample is known as the species problem. The species problem occur in the literature in many different versions: when the number of species k is the vocabulary size of Shakespeare, when k is the unknown number of components in a system, or when k is the population size. We consider Bayesian formulations of the more frequent statistical models in the literature. In all formulations the posterior distribution of k depends on the number of distinct species in the sample, the prior distribution and a quantity specified by the observational plan of the sample, as for instance size of the sample, length of time of sample observation or number of captures. We show that the posterior distribution of k is stochastically monotone in each one these parameters. These properties of monotonicity formally describe the role of data and of the prior in the Bayesian estimation of k . We show too that the Bayes estimators for k and another quantities are consistent.

Índice

1	Modelos estatísticos para amostragem de espécies	1
1.1	Introdução	1
1.2	A formulação do modelo estatístico para amostragem de espécies	3
1.3	Os modelos para amostragem de espécies	5
1.3.1	Amostragem Bernoulli-Poisson	5
1.3.2	Amostragem multinomial	6
1.3.3	Amostragem Bernoulli múltipla	8
1.3.4	Amostragem hipergeométrica	8
1.3.5	Amostragem Bernoulli-hipergeométrica	9
1.4	Uma modelagem unificadora	10
1.5	Introdução aos modelos bayesianos para amostragem de espécies	11
2	Estimativas bayesianas do número de componentes de uma superposição de processos de Poisson	16
2.1	Introdução	16
2.2	O modelo e a função de verossimilhança	19
2.3	A estrutura da priori	21
2.4	A estimação de k	24
2.5	A estimação de $U(t)$ e $V(t)$	29
2.6	Resultados assintóticos	36

3	Estimação bayesiana do número de componentes de uma superposição de processos de Poisson: generalizações e propriedades	41
3.1	Introdução	41
3.2	O modelo e a função de verossimilhança	44
3.3	A distribuição a posteriori de k	45
3.4	Propriedades da distribuição a posteriori de k	48
3.5	A estimação de $U(t)$ e $V(t)$	56
3.6	Resultados assintóticos	66
4	Estimação bayesiana do número de espécies de uma população	68
4.1	Introdução	68
4.2	O modelo estatístico e a função de verossimilhança . . .	74
4.3	A estrutura da priori	75
4.4	Características do modelo bayesiano	76
4.5	Resultados numéricos	81
4.6	Propriedades de monotonicidade estocástica da distribuição a posteriori de k	82
4.7	A estimação de $U(n)$ e $V(n)$	90
4.8	Resultados assintóticos	98
5	Estimação Bayesiana do número de espécies : amostragem por captura	104
5.1	Introdução	104
5.2	A distribuição a posteriori de k	106
5.3	Monotonicidade da distribuição a posteriori de k	108
5.4	Superposição de um número desconhecido de modelos independentes	113

6	Estimação Bayesiana do número de espécies de uma população finita	116
6.1	Introdução	116
6.2	Amostragem hipergeométrica	117
6.3	Amostragem Bernoulli-hipergeométrica	119
	Comentários finais	127
	Referências Bibliográficas	129

Capítulo 1

Modelos estatísticos para amostragem de espécies

1.1 Introdução

Consideremos uma amostra selecionada de uma população de animais constituída de várias espécies. Seja a_r o número de espécies distintas que aparecem exatamente r vezes na amostra, $r = 1, 2, \dots$. O vetor

$$a = (a_1, a_2, \dots)$$

se denomina uma amostra de espécies e Good [30] denomina a o vetor frequência das frequências, desde que a_r é a frequência da frequência r , $r = 1, 2, \dots$. Observemos que

$$\sum_{r \geq 1} r a_r$$

é o tamanho da amostra e

$$w = \sum_{r \geq 1} a_r$$

é o número de espécies distintas observadas na amostra.

Seja $\mathcal{N} = \{0, 1, 2, \dots\}$. O conjunto de todas as possíveis amostras de espécies é dado por

$$\Omega = \{(a_1, a_2, \dots) : a_r \in \mathcal{N}, r = 1, 2, \dots\}.$$

A formulação de um modelo estatístico para amostragem de espécies consiste na especificação de uma família de probabilidades sobre Ω . Usualmente estas probabilidades são indexadas por k , número de espécies na população, e por parâmetros adicionais $\theta_1, \theta_2, \dots, \theta_k$. Os parâmetros $k, \theta_1, \theta_2, \dots, \theta_k$ são desconhecidos e o problema das espécies consiste em estimar k baseado na amostra de espécies $a = (a_1, a_2, \dots) \in \Omega$.

O problema das espécies se apresenta sob muitas versões. A mais famosa é aquela em que o número de espécies é o tamanho do vocabulário de Shakespeare. Shakespeare escreveu 31534 palavras diferentes, das quais 14376 apareceram somente uma vez, 4343 apareceram exatamente duas vezes, etc.... A questão é quantas palavras diferentes Shakespeare conhecia, ou equivalentemente, dado que ele usou 31534 palavras distintas quantas palavras ele conhecia e não usou? Outras versões ocorrem, por exemplo, quando o número de espécies é o número de fontes de erros num novo sistema de programas para computador ou quando k é o tamanho de uma população. Outros exemplos podem ser encontrados em Goodman [31].

Em geral, este problema é abordado na literatura sob o ponto de vista frequentista. Na seção 1.2 formulamos os modelos estatísticos para amostragem de espécies basicamente em dois passos:

- (1) considerando-se parâmetros $k, \theta_1, \theta_2, \dots, \theta_k$, e um vetor aleatório $(X_1, X_2, \dots, X_k)$ com uma distribuição de probabilidade discreta conhecida dependente de $k, \theta_1, \theta_2, \dots, \theta_k$, e
- (2) considerando-se transformações de $(X_1, X_2, \dots, X_k)$ que induzam uma medida de probabilidade sobre Ω , conjunto de todas as pos-

síveis amostras de espécies. A família das medidas de probabilidades induzidas, indexadas por $k, \theta_1, \theta_2, \dots, \theta_k$, constitui-se num modelo estatístico para amostra de espécies.

Os modelos estatísticos mais frequentes na literatura são formulados seguindo os passos (1) e (2) acima e boa parte destes modelos são revisados por Bunge e Fitzpatrick [4]. Na seção 1.3 apresentamos estes modelos, na seção 1.4 apresentamos uma modelagem unificadora, no sentido de que os modelos usuais para amostragem de espécies são obtidos a partir de um modelo geral e na seção 1.5 introduzimos os modelos bayesianos, que estudaremos neste trabalho, explicando o que será apresentado nos capítulos subsequentes.

1.2 A formulação do modelo estatístico para amostragem de espécies

Suponhamos uma população particionada em k espécies $e_1, e_2, \dots, e_k$. Seleccionada uma amostra de tamanho fixado ou aleatório desta população e dados os parâmetros $k, \theta_1, \theta_2, \dots, \theta_k$, consideremos o vetor aleatório discreto $(X_1, X_2, \dots, X_k)$ com distribuição de probabilidade $P[\cdot | k, \theta_1, \theta_2, \dots, \theta_k]$, onde X_i é o número de elementos da espécie e_i , $i = 1, \dots, k$, observados na amostra. Seja

$$A_r = \sum_{i=1}^k I_{(X_i=r)}, \quad r = 1, 2, \dots,$$

ou seja, A_r é o número de espécies que ocorrem exatamente r vezes na amostra, $r = 1, 2, \dots$. Observemos que

$$\sum_{r \geq 1} r A_r = \sum_{r \geq 1} r \sum_{i=1}^k I_{(X_i=r)} = \sum_{i=1}^k X_i$$

é o tamanho da amostra.

$$W = \sum_{r \geq 1} A_r = \sum_{r \geq 1} \sum_{i=1}^k I_{(X_i=r)} = \sum_{i=1}^k I_{(X_i > 0)}$$

é o número de espécies distintas observadas na amostra e

$$A_0 = \sum_{i=1}^k I_{(X_i=0)} = k - W$$

é o número de espécies que não foram observadas na amostra.

No problema da amostragem de espécies, além de k e $\theta_1, \theta_2, \dots, \theta_k$ serem desconhecidos, as espécies $e_1, e_2, \dots, e_k$ não podem ser identificadas na amostra, implicando que $(X_1, X_2, \dots, X_k)$ não é observável. Contudo, a partição da amostra induzida pelas diferentes espécies representadas na amostra é observável, ou seja, as quantidades aleatórias $A_r, r = 1, 2, \dots$, são observáveis. O vetor $(A_1, A_2, \dots)$ induz a seguinte probabilidade sobre Ω : para todo $(a_1, a_2, \dots) \in \Omega$ tal que $w = \sum_{r \geq 1} a_r \leq k$,

$$P[A_1 = a_1, A_2 = a_2, \dots | k, \theta_1, \theta_2, \dots, \theta_k]$$

Lo conj. de todas as possíveis amostras de espécies. $W \leq K$

a_2 espécies distintas aparecem 2 vezes.

$$= \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} P[X_1 = x_1, \dots, X_k = x_k | k, \theta_1, \theta_2, \dots, \theta_k], \quad (1.1)$$

Lo me de indivíduos da espécie i

onde

$$[a_1, a_2, \dots; k] = \{(y_1, \dots, y_k) \in \mathcal{N}^k : \sum I_{\{r\}}(y_i) = a_r, r = 1, 2, \dots\}. \quad (1.2)$$

A cardinalidade de $[a_1, a_2, \dots; k]$ é

$$\#[a_1, a_2, \dots; k] = \binom{k}{w} \binom{w}{a_1, a_2, \dots}, \quad (1.3)$$

onde $w = \sum_{r \geq 1} a_r$ é o número de espécies distintas observadas na amostra.

Então, para todo $(a_1, a_2, \dots) \in \Omega$, a função de verossimilhança de k e $\theta_1, \theta_2, \dots, \theta_k$, dado $A_1 = a_1, A_2 = a_2, \dots$, é dada por

$$L_{(a_1, a_2, \dots)}(k, \theta_1, \theta_2, \dots, \theta_k) = \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} P[X_1 = x_1, \dots, X_k = x_k | k, \theta_1, \theta_2, \dots, \theta_k], \quad (1.4)$$

para $k = w, w + 1, \dots$, onde $w = \sum_{r \geq 1} a_r$ e $(\theta_1, \theta_2, \dots, \theta_k) \in \Theta_k$, onde Θ_k é um espaço paramétrico de dimensão k .

1.3 Os modelos para amostragem de espécies

Como vimos na seção anterior, os modelos estatísticos para amostragem de espécies são construídos a partir da distribuição das frequências das frequências de um vetor aleatório $(X_1, X_2, \dots, X_k)$. Nesta seção vamos apresentar modelos para amostragem de espécies classificados segundo a distribuição de $(X_1, X_2, \dots, X_k)$. Estes modelos foram abordados na literatura sob o ponto de vista frequentista e, sob esta abordagem, muitas tentativas foram feitas para se resolver o problema das espécies. Não se tem indicações de que as soluções obtidas sejam satisfatórias, (Bunge e Fitzpatrick [4]).

1.3.1 Amostragem Bernoulli-Poisson

Dados $k \in \mathcal{N}^* = \mathcal{N} - \{0\}$; $\lambda_1, \lambda_2, \dots, \lambda_k$ tais que $\lambda_i > 0$, $i = 1, 2, \dots, k$; q tal que $0 < q \leq 1$ e $t > 0$ sejam $X_1, X_2, \dots, X_k$ variáveis aleatórias independentes com distribuição de Poisson com taxas $q\lambda_1 t, q\lambda_2 t, \dots, q\lambda_k t$, respectivamente. Então, a probabilidade induzida pelo vetor $(A_1, A_2, \dots)$ é dada por

$$P[A_1 = a_1, A_2 = a_2, \dots | k, q, \lambda_1, \lambda_2, \dots, \lambda_k]$$

$$= \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots, k]} \prod_{i=1}^k \frac{e^{-q\lambda_i t} (q\lambda_i t)^{x_i}}{x_i!}, \quad (1.5)$$

para todo $(a_1, a_2, \dots) \in \Omega$ tal que $\sum_{r \geq 1} a_r \leq k$.

Observemos que, neste caso, o tamanho da amostra $n = \sum_{i=1}^k X_i$ é aleatório. Este modelo foi utilizado por Ross [57], [58]. Ele supõe que um novo sistema de programas para computador tem um número desconhecido k de fontes de erros $e_1, e_2, \dots, e_k$. A fonte e_i causa erros de acordo com um processo de Poisson com taxa desconhecida λ_i , $i = 1, 2, \dots, k$, e estes k processos são independentes. Ele supõe também que cada erro é independentemente detectado com probabilidade conhecida q e, após acionar o sistema durante um intervalo de tempo de duração t , todos os erros detectados são removidos do sistema. Seu interesse é estimar a taxa de falha restante do sistema revisado.

Nayak [52],[55], considerando o caso $q = 1$, estuda a estimação de k e a estimação da taxa de falha restante do sistema revisado. Se não houver restrição alguma no espaço paramétrico não existe um estimador não viciado para k e o estimador de máxima verossimilhança pode ser infinito com probabilidade positiva. Mesmo sob a restrição de que todas as taxas sejam iguais, não existe um estimador não viciado para k e o estimador de máxima verossimilhança é infinito se e somente se $w = \sum_{r=1}^n a_r = n$. Veja Nayak [50],[51].

Este modelo, no caso $q = t = 1$, foi originalmente proposto por Fisher, Corbet e Willians [25] para estudar a relação entre o número de espécies distintas e o tamanho da amostra.

1.3.2 Amostragem multinomial

Dados n e $k \in \mathcal{N}^*$ e $p_1, p_2, \dots, p_k$ tais que $p_i > 0$, $i = 1, 2, \dots, k$ e $\sum_{i=1}^k p_i = 1$, seja $(X_1, X_2, \dots, X_k)$ um vetor aleatório k -dimensional com distribuição multinomial com parâmetros n e $(p_1, p_2, \dots, p_k)$. En-

tão, a probabilidade induzida pelo vetor $(A_1, A_2, \dots)$ é dada por

$$P[A_1 = a_1, A_2 = a_2, \dots | k, p_1, p_2, \dots, p_k] \\ = \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} \binom{n}{x_1, \dots, x_k} p_1^{x_1} \dots p_k^{x_k}, \quad (1.6)$$

para todo $(a_1, a_2, \dots) \in \Omega$ tal que $\sum_{r \geq 1} a_r \leq k$. Este é o modelo mais frequente na literatura para se resolver o problema das espécies.

Se $p_1 = p_2 = \dots = p_k = 1/k$, o modelo se reduz a

$$P[A_1 = a_1, A_2 = a_2, \dots | k] \\ = \binom{k}{w} \binom{w}{a_1, a_2, \dots} \binom{n}{\underbrace{1, \dots, 1}_{a_1}, \underbrace{2, \dots, 2}_{a_2}, \dots} \left(\frac{1}{k}\right)^n. \quad (1.7)$$

Neste caso os procedimentos tradicionais de estimação (máxima verossimilhança e não viciado de variância mínima) são estudados por vários autores, incluindo Lewontin e Prout [45], Darroch [16], Harris [33], Johnson e Kotz [38, pag. 136-139] e Holst [36]. Estes estimadores dependem da amostra de espécies somente através de w , número de espécies distintas observadas na amostra. Isto pode ser visto em (1.7).

O caso não equiprovável é muito mais interessante nas aplicações, mas muito mais difícil de ser trabalhado. Não existe um estimador não viciado para k se nenhuma restrição for colocada no modelo (Engen[21, pag. 28]), e o estimador de máxima verossimilhança é infinito para quaisquer resultados amostrais (Boender e Rinnooy Kan [2]). Chao [7] e Chao e Lee [9] obtêm estimadores para k que dependem das frequências das frequências $a_1, a_2, \dots$. Há indicações empíricas de que estes estimadores se comportam melhor do que os outros, mas não existe uma justificativa formal para este melhor comportamento (Bunge e Fitzpatrick [4]).

1.3.3 Amostragem Bernoulli múltipla

Dados $k \in \mathcal{N}^*$, $q_1, q_2, \dots, q_k$ tais que $0 < q_i \leq 1$, $i = 1, 2, \dots, k$, e $m \in \mathcal{N}^*$ sejam $X_1, X_2, \dots, X_k$ variáveis aleatórias independentes com distribuição Binomial com parâmetros $(m, q_1), (m, q_2), \dots, (m, q_k)$, respectivamente. Então, a probabilidade induzida por $(A_1, A_2, \dots)$ é

$$P[A_1 = a_1, A_2 = a_2, \dots | k, q_1, q_2, \dots, q_k] \\ = \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots, k]} \prod_{i=1}^k \binom{m}{x_i} q_i^{x_i} (1 - q_i)^{m-x_i}, \quad (1.8)$$

para todo $(a_1, a_2, \dots) \in \Omega$ tal que $\sum_{r \geq 1} r a_r \leq km$ e $\sum_{r \geq 1} a_r \leq k$. Notemos que o tamanho da amostra $n = \sum_{i=1}^k X_i$ é aleatório.

Este modelo foi utilizado por Burnhan e Overton [5]. Eles supõem que uma população infinita particionada em k classes $e_1, e_2, \dots, e_k$, $k \in \mathcal{N}^*$, é observada em m ocasiões diferentes. Em cada ocasião um elemento da classe e_i é capturado com probabilidade q_i ou não capturado com probabilidade $1 - q_i$, e X_i é o número de elementos da classe e_i capturados nas m ocasiões, $i = 1, 2, \dots, k$. Supondo que, dado $k, q_1, q_2, \dots, q_k$ são independentes e identicamente distribuídos com distribuição F , F desconhecido, eles obtêm um estimador para k pelo método jackknife. Este modelo também foi considerado por Chao [8] e Yip [67]. Não se tem ainda resultados conclusivos acerca dos estimadores obtidos por estes autores.

1.3.4 Amostragem hipergeométrica

Dados k, N e $n \in \mathcal{N}^*$ tais que $N \geq n$ e $N \geq k$ e dado $(N_1, N_2, \dots, N_k)$ tal que $N_i \in \mathcal{N}$, $i = 1, 2, \dots, k$ e $\sum_{i=1}^k N_i = N$, seja $(X_1, X_2, \dots, X_k)$ um vetor aleatório k -dimensional com distribuição hipergeométrica com parâmetros N, n e $(N_1, N_2, \dots, N_k)$. Então, a probabilidade induzida por $(A_1, A_2, \dots)$ é dada por

$$\begin{aligned}
P[A_1 = a_1, A_2 = a_2, \dots | k, N_1, N_2, \dots, N_k] \\
= \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} \frac{\prod_{i=1}^k \binom{N_i}{x_i}}{\binom{N}{n}}, \tag{1.9}
\end{aligned}$$

para todo $(a_1, a_2, \dots) \in \Omega$ tal que $\sum_{r \geq 1} r a_r = n$ e $\sum_{r \geq 1} a_r \leq k$.

Este modelo foi utilizado por Goodman [31], Chapman [11], Korwar [40] e Shlosser [60]. Goodman mostra que se nenhuma restrição for colocada no modelo, não existe um estimador não viciado para k . Contudo, se $n \geq \max_{1 \leq i \leq k} N_i$ existe um único estimador não viciado para k e, por ser único, é de variância uniformemente mínima, ainda que em muitos casos esta variância mínima seja grande, Engen [21].

1.3.5 Amostragem Bernoulli-hipergeométrica

Dados $k, N \in \mathcal{N}^*$ com $N \geq k$; $N_1, N_2, \dots, N_k$ tais que $N_i \in \mathcal{N}^*$, $i = 1, 2, \dots, k$ e $\sum_{i=1}^k N_i = N$, e q , $0 < q \leq 1$, sejam $X_1, X_2, \dots, X_k$ variáveis aleatórias independentes com distribuição Binomial com parâmetros $(N_1, q), (N_2, q), \dots, (N_k, q)$, respectivamente. A probabilidade induzida por $(A_1, A_2, \dots)$ é

$$\begin{aligned}
P[A_1 = a_1, A_2 = a_2, \dots | k, q, N_1, N_2, \dots, N_k] \\
= \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} q^n (1 - q)^{N-n} \prod_{i=1}^k \binom{N_i}{x_i} \tag{1.10}
\end{aligned}$$

onde $n = \sum_{r \geq 1} r a_r$, para todo $(a_1, a_2, \dots) \in \Omega$ tal que $\sum_{r \geq 1} r a_r \leq N$ e $\sum_{r \geq 1} a_r \leq k$.

Neste caso o tamanho da amostra é $\sum_{i=1}^k X_i$ que é uma variável aleatória com distribuição Binomial(N, q). Supondo N e q conhecidos Goodman [31] derivou um estimador não viciado para k mas, como

observa Frank [24], a variância deste estimador é muito grande. Outros pesquisadores, como por exemplo, Harwit [34], Esty [22], Wani e Lo [64],[65], Esty [22],[23], utilizaram este modelo.

1.4 Uma modelagem unificadora

Nesta seção vamos apresentar um modelo estatístico para amostragem de espécies, a partir do qual podem ser derivados os modelos apresentados na seção anterior.

Sejam $k \in \mathcal{N}^*$; $\lambda_1, \lambda_2, \dots, \lambda_k$, tais que $\lambda_i > 0$, $i = 1, 2, \dots, k$; $q_1, q_2, \dots, q_k$, tais que $0 < q_i \leq 1$, $i = 1, 2, \dots, k$, e $t > 0$. Consideremos as variáveis aleatórias $N_1, N_2, \dots, N_k$, e $X_1, X_2, \dots, X_k$, satisfazendo as seguintes suposições:

- (suposição 1) as variáveis aleatórias $N_1, N_2, \dots, N_k$ são independentes com distribuição de Poisson com parâmetros $\lambda_1 t, \lambda_2 t, \dots, \lambda_k t$, respectivamente, e
- (suposição 2) dados $N_1, N_2, \dots, N_k$, as variáveis aleatórias $X_1, X_2, \dots, X_k$ são independentes com distribuição Binomial com parâmetros $(N_1, q_1), (N_2, q_2), \dots, (N_k, q_k)$, respectivamente.

Pelas suposições 1 e 2 segue que, dados $k; \lambda_1, \lambda_2, \dots, \lambda_k; q_1, q_2, \dots, q_k$, e $t > 0$, as variáveis aleatórias $X_1, X_2, \dots, X_k$ são independentes com distribuição de Poisson com taxas $q_1 \lambda_1 t, q_2 \lambda_2 t, \dots, q_k \lambda_k t$, respectivamente. Recorramos à formulação geral do problema na seção 1.2; então, o vetor $(A_1, A_2, \dots)$ induz a seguinte probabilidade sobre $\Omega = \{(a_1, a_2, \dots) : a_r \in \mathcal{N}, r = 1, 2, \dots\}$: para todo $(a_1, a_2, \dots) \in \Omega$ tal que $\sum_{r \geq 1} a_r \leq k$,

$$P[A_1 = a_1, A_2 = a_2, \dots | k; \lambda_1, \lambda_2, \dots, \lambda_k; q_1, q_2, \dots, q_k; t]$$

$$= \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots, k]} \prod_{i=1}^k \frac{e^{-q_i \lambda_i t} (q_i \lambda_i t)^{x_i}}{x_i!}. \quad (1.11)$$

Se em (1.11) fizermos $q_1 = q_2 = \dots = q_k = q$, $0 < q \leq 1$, obteremos (1.5) e teremos o modelo amostragem Bernoulli-Poisson; como a distribuição de $(X_1, X_2, \dots, X_k)$, dado que $\sum_i X_i = n$, é multinomial com parâmetros n e $p_1, p_2, \dots, p_k$, onde $p_i = \frac{q_i \lambda_i}{\sum_{j=1}^k q_j \lambda_j}$, $i = 1, 2, \dots, k$, então, se considerarmos a distribuição de $(A_1, A_2, \dots)$, dado que $\sum_{i=1}^k X_i = n$, ela será a soma de termos multinomiais como em (1.6) e teremos o modelo amostragem multinomial; como a distribuição de $(X_1, X_2, \dots, X_k)$, dados $N_1, N_2, \dots, N_k$ e $\sum_{i=1}^k X_i = n$, é hipergeométrica com parâmetros $N = \sum_{i=1}^k N_i$, n e $N_1, N_2, \dots, N_k$, então se considerarmos a distribuição de $(A_1, A_2, \dots)$, dados $N_1, N_2, \dots, N_k$ e $\sum_{i=1}^k X_i = n$, ela será a soma de termos da distribuição hipergeométrica como em (1.9) e teremos o modelo amostragem hipergeométrica; como, pela suposição 2, dados $N_1, N_2, \dots, N_k$ as variáveis aleatórias $X_1, X_2, \dots, X_k$ são independentes com distribuição Binomial com parâmetros $(N_1, q_1), (N_2, q_2), \dots, (N_k, q_k)$, respectivamente, então se considerarmos a distribuição de $(A_1, A_2, \dots)$, dados $N_1, N_2, \dots, N_k$, com a restrição $q_1 = q_2 = \dots = q_k = q$, obteremos (1.10) e teremos o modelo amostragem Bernoulli-hipergeométrica. Finalmente, se considerarmos a distribuição de $(A_1, A_2, \dots)$, dado $N_1 = N_2 = \dots = N_k = m$, obteremos (1.8) e teremos o modelo amostragem Bernoulli múltipla.

1.5 Introdução aos modelos bayesianos para amostragem de espécies

A literatura sobre o problema das espécies é proeminentemente frequentista, enquanto as abordagens bayesianas do problema não são tão frequentes. Fazemos nesta seção uma breve revisão da literatura

bayesiana sobre este tema. apresentamos uma idéia geral dos resultados de nossa pesquisa sobre o assunto e descrevemos a forma em que estes resultados serão apresentados nesta tese.

Vimos brevemente na seção 1.3 que a literatura frequentista não tem ainda uma solução satisfatória para o problema das espécies. Por outro lado, as poucas soluções bayesianas apresentadas na literatura não foram formalmente analisadas. Somente dois modelos estatísticos foram bayesianamente formulados: o modelo por amostragem multinomial (modelo(1.6)) e o modelo por amostragem hipergeométrica (modelo(1.9)). Dificuldades computacionais foram encontradas e a distribuição a posteriori de k , com exceção de casos muito particulares, não tem uma forma explícita. Hill [35] considera a amostragem hipergeométrica de uma população finita de tamanho N (modelo (1.9)), supõe que, dado k , as k frequências absolutas de cada classe da população, $N_1, N_2, \dots, N_k$, tem distribuição uniforme no conjunto $\{(f_1, f_2, \dots, f_k) : \sum_{i=1}^k f_i = N, f_i \geq 1, i = 1, 2, \dots, k\}$ e que k tem distribuição Binomial Negativa truncada, obtendo a média e a variância a posteriori de k . Lewins e Joanes [44] consideram o modelo de amostragem multinomial (modelo(1.6)), supõem que, dado k , $(p_1, p_2, \dots, p_k)$ tem distribuição Dirichlet simétrica com parâmetro α e que k tem distribuição Binomial Negativa truncada. Boender e Rinnooy Kan [3] consideram também a amostragem multinomial, supõem que, dado k , as proporções populacionais têm distribuição Dirichlet simétrica com parâmetro α e consideram uma priori imprópria uniforme em $\mathcal{N}^*$ para k . Lewins e Joanes [44] e Boender e Rinnooy Kan [3] verificam numericamente que a média e a moda da distribuição a posteriori são muito sensíveis ao valor de α . Leite e Bolfarine [42] consideram a amostragem multinomial no caso equiprovável $p_1 = p_2 = \dots = p_k = \frac{1}{k}$ (modelo(1.7)), supõem uma priori qualquer para k e estudam analiticamente propriedades monotônicas da média a posteriori de k .

Neste trabalho consideramos formulações bayesianas dos modelos estatísticos descritos na seção 1.3. Em cada um destes modelos estatísticos, a amostra de espécies de uma população com k espécies é modelada pela distribuição das frequências das frequências $A_1, A_2, \dots$ de um vetor aleatório $(X_1, X_2, \dots, X_k)$ com distribuição dependendo de um conjunto de parâmetros (*nuisance*) $\theta_1, \theta_2, \dots, \theta_k$. Os parâmetros desconhecidos são $k, \theta_1, \theta_2, \dots, \theta_k$ e os dados são os valores observados de $A_1, A_2, \dots$. Na falta de uma distribuição conjugada para $(k, \theta_1, \theta_2, \dots, \theta_k)$ com relação a $A_1, A_2, \dots$, vamos sempre supor que, dado $k, (\theta_1, \theta_2, \dots, \theta_k)$ tem uma distribuição conjugada com relação a $(X_1, X_2, \dots, X_k)$; k terá sempre uma distribuição arbitrária qualquer. Em cada uma das formulações bayesianas, a distribuição a posteriori de k depende da priori sobre os parâmetros *nuisance*, do tempo de observação da amostra ou o número de capturas na amostra e dos dados, através do número de espécies observadas na amostra. Nosso principal resultado consiste em explicar a função de cada um destes parâmetros na estimação de k . Mostramos que a distribuição a posteriori de k é estocasticamente monótona em cada um destes parâmetros.

No capítulo 2 consideramos a formulação bayesiana da amostragem Bernoulli-Poisson (modelo(1.5)). Supomos que as taxas dos processos são independentes e identicamente distribuídas com distribuição Gama. Tais suposições são convenientes, pois para uma grande classe de priors para k as posteriors para k podem ser obtidas explicitamente. As estimativas para as taxas dos processos não observados também são obtidas explicitamente. Mostramos que todos os estimadores são consistentes.

No capítulo 3 consideramos a formulação bayesiana do modelo unificador de amostragem (modelo (1.11)). Este modelo bayesiano corresponde àquele considerado no capítulo 2, supondo que as taxas dos processos detectados são independentes e identicamente distribuídas

como uma distribuição F qualquer. Para uma grande classe de priors, determinamos explicitamente posterioris para k e para as taxas dos processos não observados na amostra. A distribuição a posteriori de k depende da distribuição das taxas dos processos, F , do tempo de observação do sistema e do número de processos observados. Mostramos que a distribuição a posteriori de k é decrescente em F , ou seja, quanto maiores as taxas dos erros detectados menor será a estimativa de Bayes para k . Mostramos também que a distribuição a posteriori de k é crescente no número de processos observados e decrescente no tempo de observação do sistema.

No capítulo 4 consideramos a formulação bayesiana da amostragem multinomial (modelo(1.6)). Supomos uma priori arbitrária para k e que, dado k , as proporções tem distribuição Dirichlet simétrica com parâmetro α . A distribuição a posteriori de k depende de α , do tamanho da amostra, n , e do número de espécies distintas observadas na amostra, w . Mostramos que a distribuição a posteriori de k é estocasticamente monótona em cada um destes parâmetros. Mostramos que a distribuição a posteriori de k é estocasticamente decrescente em α e, com isto, temos uma justificativa formal para o papel de α como a abundância relativa das espécies, com α pequeno implicando grandes estimativas de Bayes para k . Mostramos também que a distribuição a posteriori de k é estocasticamente crescente em w e decrescente em n . Obtemos expressões para a média e a variância a posteriori de k , para a média e a variância a posteriori da probabilidade condicional de se descobrir uma nova espécie, em uma observação adicional, e para a correspondente probabilidade incondicional de se descobrir uma nova espécie. Verificamos que todos estes estimadores são consistentes.

No capítulo 5 consideramos a formulação bayesiana do modelo (1.8), modelo proposto por Burnham e Overton [5] para estimar k , número de espécies de uma população onde as probabilidades de captura variam

entre as espécies. Neste modelo a função de verossimilhança de k , dados os dados, produz sempre posterioris que podem ser obtidas explicitamente para uma grande classe de prioris. Verificamos que a distribuição a posteriori de k é estocasticamente decrescente nas probabilidades de captura, crescente no número de espécies capturadas e decrescente no número ocasiões de capturas. Estudamos também o comportamento assintótico dos momentos a posteriori de k .

No capítulo 6 consideramos a formulação bayesiana dos modelos de população finita (1.9) e (1.10). Enunciamos várias propriedades de ordenação estocástica da distribuição a posteriori de k .

Capítulo 2

Estimativas bayesianas do número de componentes de uma superposição de processos de Poisson

2.1 Introdução

Consideremos uma superposição de um número desconhecido, k , de processos de Poisson independentes e homogêneos, $\{X_i(t), t \geq 0\}$, $i = 1, 2, \dots, k$, com taxas desconhecidas $\lambda_1, \lambda_2, \dots, \lambda_k$. Suponhamos que, toda vez que um evento ocorre podemos identificar o processo que o gerou, mas não podemos identificar o índice que o enumera como um componente do sistema. Por exemplo, um número desconhecido de fontes (como erros num grande sistema de programas de computador, objetos celestiais, ou usuários de um serviço específico) transmite sinais de acordo com um processo de Poisson homogêneo, onde a fonte de cada sinal é identificada quando ela ocorre e não por um índice que pudesse estar estampado na fonte. Estamos considerando que os processos $\{X_i(t), t \geq 0\}$, $i = 1, 2, \dots, k$, não são observáveis, mas a partição induzida pelos diferentes processos identificados na amostra é

observável. Seja $t \geq 0$ e definamos

$$A_j(t) = \sum_{i=1}^k I_{(X_i(t)=j)}, \quad j \in \mathcal{N}^*,$$

$$W_t = \sum_{j \geq 1} A_j(t)$$

e

$$N_t = \sum_{j \geq 1} j A_j(t).$$

Observemos que para todo $j \in \mathcal{N}^*$,

$A_j(t)$ é o número de processos observados exatamente j vezes durante o intervalo de tempo $[0, t]$, $t \geq 0$, com $A_j(0) = 0$;

$A_0(t)$ é o número de processos não observados durante o intervalo de tempo $[0, t]$, $t \geq 0$, com $A_0(0) = k$;

$W_t = \sum_{i=1}^k I_{(X_i(t)>0)}$ é o número total de processos distintos observados durante o intervalo de tempo $[0, t]$, $t \geq 0$, com $W_0 = 0$;

$N_t = \sum_{i=1}^k X_i(t)$ é o número total de ocorrências dos k processos durante o intervalo de tempo $[0, t]$, $t \geq 0$;

$$\sum_{j \geq 0} A_j(t) = k \text{ e}$$

$$A_0(t) = k - W_t, \text{ para todo } t \geq 0.$$

Se o sistema for observado durante um intervalo de tempo de duração t , $t > 0$, $A_j(t)$, $j \in \mathcal{N}^*$, são quantidades observáveis durante o intervalo $[0, t]$. Como $X_1(t), X_2(t), \dots, X_k(t)$ não são observáveis, $A_0(t) = k - W_t$ também não é observável. Após observarmos o sistema durante um intervalo de tempo fixado $[0, t]$, $t > 0$, queremos estimar

k , número de processos.

$U(t) = \sum_{i=1}^k \lambda_i I_{(X_i(t)=0)}$, taxa total dos processos não observados durante o intervalo de tempo $[0, t]$ e

$V(t) = E[U(t)] = \sum_{i=1}^k \lambda_i e^{-\lambda_i t}$, taxa total esperada dos processos não observados durante o intervalo de tempo $[0, t]$.

A estimação de k é um problema de difícil solução (Nayak [52]). Um estimador não viciado para k não existe, mesmo no caso especial em que os λ_i são iguais. Neste caso o estimador de máxima verossimilhança de k é infinito se $W_t = N_t$ (Nayak [50], [51]). A estimação de $V(t)$ é um problema mais tratável. Nayak [51] e [52] prova que $A_1(t)/t$ é o estimador não viciado de variância uniformemente mínima de $V(t)$. Já a estimação de $U(t)$ é um problema para o qual os métodos usuais (por exemplo, o método de máxima verossimilhança) não se aplicam (veja Nayak [55]). Ross [57] observa que $A_1(t)/t$ pode ser considerado como um estimador de $U(t)$ desde que

$$E\left[\frac{A_1(t)}{t}\right] = E[U(t)],$$

e mostra que

$$\text{Cov}\left[\frac{A_1(t)}{t}, U(t)\right] < 0.$$

Nayak [55] observa que esta covariância negativa é uma propriedade indesejável, pois ela aumenta o erro quadrático médio $E\left[\frac{A_1(t)}{t} - U(t)\right]^2$, mostra que todos os estimadores não viciados de $V(t)$ são negativamente correlacionados com $U(t)$ e que um estimador de erro quadrático médio uniformemente mínimo de $U(t)$ não existe.

Neste capítulo propomos um procedimento bayesiano para resolver este problema. De acordo com a ótica bayesiana, todas as quantidades do sistema são de dois tipos: aquelas que são conhecidas e aquelas que

são desconhecidas pelo estatístico. Se o sistema for observado durante um intervalo de tempo fixado $[0, t]$, $t > 0$, as quantidades conhecidas são descritas pelos valores $(a_1, a_2, \dots)$ de $(A_1(t), A_2(t), \dots)$ e as quantidades desconhecidas $k, \lambda_1, \lambda_2, \dots, \lambda_k$ e $X_1(t), X_2(t), \dots, X_k(t)$ são descritas por uma distribuição de probabilidade conjunta. Como as quantidades que queremos estimar são k , $U(t)$ e $V(t)$ então nosso interesse, em primeiro lugar, é obter as distribuições de probabilidade de k , $U(t)$ e $V(t)$, dado $(a_1, a_2, \dots)$.

Na seção 2.2 apresentamos o modelo de amostragem e a função de verossimilhança. Na seção 2.3 descrevemos a estrutura da priori para $k, \lambda_1, \lambda_2, \dots, \lambda_k$. Dado k , supomos que $\lambda_1, \lambda_2, \dots, \lambda_k$ são variáveis aleatórias independentes e identicamente distribuídas com distribuição Gama e que k tem uma distribuição a priori qualquer. Na seção 2.4, obtemos várias posteriores explícitas para k . Na seção 2.5 obtemos as distribuições, as médias e as variâncias a posteriori de $U(t)$ e $V(t)$. Na seção 2.6 estudamos as propriedades assintóticas das estimativas de Bayes de k , $U(t)$ e $V(t)$. Verificamos que, quando $t \rightarrow \infty$, a média a posteriori de k converge quase certamente para $\tau(k) = \inf\{n \in \mathcal{N} : n \geq k \text{ e } \pi(n) > 0\}$, onde π é uma priori para k ; verificamos também que as médias e as variâncias a posteriori de $U(t)$ e $V(t)$ convergem quase certamente para zero.

2.2 O modelo e a função de verossimilhança

Suponhamos que o sistema seja observado durante o intervalo de tempo $[0, t]$, $t > 0$, e sejam $k \in \mathcal{N}^*$ e $\lambda_j \in (0, \infty)$, $j = 1, 2, \dots, k$. Pelas suposições iniciais temos que

$$P[X_1(t) = x_1, \dots, X_k(t) = x_k | k, \lambda_1, \dots, \lambda_k] = \prod_{i=1}^k e^{-\lambda_i t} \frac{(\lambda_i t)^{x_i}}{x_i!}, \quad (2.1)$$

para quaisquer $x_1, \dots, x_k \in \mathcal{N}$. Assim, a distribuição das quantidades observáveis $(A_j(t), j \in \mathcal{N}^*)$ é uma soma de termos como em (2.1) e pode ser escrita, para quaisquer $a_1, a_2, \dots \in \mathcal{N}$ tal que $w = \sum_{i \geq 1} a_i \leq k$, como

$$\begin{aligned} P[A_1(t) = a_1, A_2(t) = a_2, \dots | k, \lambda_1, \dots, \lambda_k] \\ = \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} \prod_{i=1}^k e^{-\lambda_i t} \frac{(\lambda_i t)^{x_i}}{x_i!}, \end{aligned} \quad (2.2)$$

onde

$$\begin{aligned} [a_1, a_2, \dots; k] \\ = \{(y_1, \dots, y_k) : y_1, \dots, y_k \in \mathcal{N}; \sum_{i=1}^k I_{\{j\}}(y_i) = a_j, j \geq 1\}. \end{aligned}$$

Observemos que o número de parcelas na soma (2.2) é dada por

$$\#[a_1, a_2, \dots; k] = \binom{k}{a_0, a_1, \dots},$$

onde $a_0 = k - w$. Esta formulação é considerada por Nayak [52]. Para quaisquer $a_1, a_2, \dots \in \mathcal{N}$ tal que $w = \sum_{j \geq 1} a_j < \infty$, a função de verossimilhança de $(k, \lambda_1, \dots, \lambda_k)$ dado $(a_1, a_2, \dots)$ é

$$\begin{aligned} L_{(a_1, a_2, \dots)}(k, \lambda_1, \dots, \lambda_k) \\ = \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} \prod_{i=1}^k e^{-\lambda_i t} \frac{(\lambda_i t)^{x_i}}{x_i!} I_{(0, \infty)}(\lambda_i) I_{\{w, w+1, \dots\}}(k). \end{aligned} \quad (2.3)$$

As quantidades que queremos estimar são k , número de componentes do sistema, $U(t)$, taxa total dos processos não observados e $V(t)$, taxa total esperada dos processos não observados.

2.3 A estrutura da priori

Seja π uma função de probabilidade a priori para k definida no conjunto $\mathcal{N}^*$. Suponhamos que dado k , $k \in \mathcal{N}^*$, $\lambda_1, \lambda_2, \dots, \lambda_k$ sejam variáveis aleatórias independentes e identicamente distribuídas com distribuição Gama com parâmetros α e β , $\alpha > 0$, $\beta > 0$. Então, a distribuição a priori conjunta de $(\lambda_1, \lambda_2, \dots, \lambda_k, k)$ é

$$\begin{aligned} \pi(\lambda_1, \lambda_2, \dots, \lambda_k, k) &= \pi(\lambda_1, \lambda_2, \dots, \lambda_k | k) \pi(k) \\ &= \frac{\beta^{k\alpha}}{\Gamma^k(\alpha)} \prod_{i=1}^k \lambda_i^{\alpha-1} e^{-\beta\lambda_i} I_{(0,\infty)}(\lambda_i) \pi(k) I_{\mathcal{N}^*}(k). \end{aligned} \quad (2.4)$$

Proposição 2.3.1 *Com as suposições acima temos as seguintes propriedades:*

- (i) dados k e $X_1(t) = x_1, X_2(t) = x_2, \dots, X_k(t) = x_k$, $\lambda_1, \lambda_2, \dots, \lambda_k$ são variáveis aleatórias independentes com distribuição Gama com parâmetros $\alpha + x_i$ e $\beta + t$, $i = 1, 2, \dots, k$;
- (ii) dados k e $X_1(t) = x_1, X_2(t) = x_2, \dots, X_k(t) = x_k$ tal que $w = \sum_{i=1}^k I_{\mathcal{N}^*}(x_i) < k$, $U(t) = \sum_{i=1}^k \lambda_i I_{(X_i(t)=0)}$ é uma variável aleatória com distribuição Gama com parâmetros $\alpha(k - w)$ e $\beta + t$;
- (iii) dado k , $X_1(t), X_2(t), \dots, X_k(t)$ são independentes e identicamente distribuídas como uma Binomial Negativa com parâmetros α e $\beta/(\beta + t)$, ou seja,

$$P[X_i(t) = x | k] = \frac{\Gamma(\alpha + x)}{\Gamma(\alpha) x!} \left(\frac{\beta}{\beta + t}\right)^\alpha \left(\frac{t}{\beta + t}\right)^x I_{\mathcal{N}}(x),$$

$i = 1, 2, \dots, k$, e assim

$$\begin{aligned} P[X_1(t) = x_1, X_2(t) = x_2, \dots, X_k(t) = x_k | k] & \\ &= \left(\frac{\beta}{\beta + t}\right)^{k\alpha} \prod_{i=1}^k \frac{\Gamma(\alpha + x_i)}{\Gamma(\alpha) x_i!} \left(\frac{t}{\beta + t}\right)^{x_i} I_{\mathcal{N}}(x_i); \end{aligned} \quad (2.5)$$

(iv) dado k , W_t é uma variável aleatória Binomial com parâmetros k e $1 - (\beta/(\beta + t))^\alpha$. assim,

$$E[W_t|k] = k(1 - (\frac{\beta}{\beta + t})^\alpha)$$

e

$$Var[W_t|k] = k(1 - (\frac{\beta}{\beta + t})^\alpha)(\frac{\beta}{\beta + t})^\alpha;$$

(v) a distribuição marginal de W_t é

$$P[W_t = w] = C(\alpha, \beta, t, w) (\frac{1}{(\frac{\beta}{\beta+t})^\alpha} - 1)^w I_N(w),$$

onde

$$C(\alpha, \beta, t, w) = \sum_{k \geq w} \binom{k}{w} (\frac{\beta}{\beta + t})^{k\alpha} \pi(k);$$

(vi) dado k , a distribuição de $(A_1(t), A_2(t), \dots)$ é, para quaisquer $a_1, a_2, \dots \in \mathcal{N}$ tal que $w = \sum_{j \geq 1} a_j \leq k$,

$$\begin{aligned} & P[A_1(t) = a_1, A_2(t) = a_2, \dots | k] \\ &= \binom{k}{a_0, a_1, \dots} (\frac{t}{\beta + t})^{a_t} (\frac{\beta}{\beta + t})^{k\alpha} \prod_{i \geq 1} (\frac{\Gamma(\alpha + i)}{\Gamma(\alpha) i!})^{a_i}, \end{aligned}$$

onde $a_0 = k - w$ e $a_t = \sum_{j \geq 1} j a_j$.

Prova. (i) segue das propriedades da distribuição Gama; (ii) segue da propriedade reprodutiva da distribuição Gama; (iii) segue do fato de que a distribuição Binomial Negativa pode ser obtida misturando-se uma distribuição Poisson com uma distribuição Gama; (iv) segue do fato de que dado k , $I_{(X_i(t) > 0)}$, $i = 1, 2, \dots, k$, são variáveis aleatórias Bernoulli condicionalmente independentes com parâmetros

$$P[X_i(t) > 0 | k] = 1 - P[X_i(t) = 0 | k] = 1 - (\frac{\beta}{\beta + t})^\alpha.$$

A propriedade (v) segue imediatamente da propriedade (iv) e finalmente, para quaisquer $a_1, a_2, \dots \in \mathcal{N}$ tal que $w = \sum_{j \geq 1} a_j \leq k$,

$$\begin{aligned} & P[A_1(t) = a_1, A_2(t) = a_2, \dots | k] \\ &= \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} P[X_1(t) = x_1, \dots, X_k(t) = x_k | k]. \end{aligned}$$

Assim, segue de (2.2) que

$$\begin{aligned} & P[A_1(t) = a_1, A_2(t) = a_2, \dots | k] \\ &= \left(\frac{\beta}{\beta + t}\right)^{k\alpha} \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} \prod_{i \geq 1} \left(\frac{\Gamma(\alpha + i)}{\Gamma(\alpha) i!}\right)^{a_i} \left(\frac{t}{\beta + t}\right)^{a_i} \\ &= \binom{k}{a_0, a_1, \dots} \left(\frac{t}{\beta + t}\right)^{a_t} \left(\frac{\beta}{\beta + t}\right)^{k\alpha} \prod_{i \geq 1} \left(\frac{\Gamma(\alpha + i)}{\Gamma(\alpha) i!}\right)^{a_i}, \end{aligned}$$

onde $a_0 = k - w$ e $a_t = \sum_{j \geq 1} j a_j$, provando (vi). ◇

Para toda função de probabilidade a priori π para k e $x \in \mathcal{N}^*$, seja

$$A_{\pi, x} = \{n \in \mathcal{N}^* : n \geq x \text{ e } \pi(n) > 0\}.$$

Pela propriedade (vi), para quaisquer $a_1, a_2, \dots \in \mathcal{N}$ tal que $w = \sum_{j \geq 1} a_j < \infty$, a função de verossimilhança de k , dado $(a_1, a_2, \dots)$, é proporcional a

$$K_w(k) = \binom{k}{w} \left[\left(\frac{\beta}{\beta + t}\right)^\alpha\right]^k, \quad (2.6)$$

ou seja, $K_w(k)$ é o *kernel* da função de verossimilhança de k , dado $(a_1, a_2, \dots)$. Se $A_{\pi, w} \neq \emptyset$, então a função de probabilidade a posteriori de k , dado $(a_1, a_2, \dots)$, está bem definida, depende dos dados somente através de $w = \sum_{j \geq 1} a_j$ e pode ser escrita como

$$\pi(k | \alpha, \beta, t, w) = \pi(k | \alpha, \beta, t, W_t = w)$$

$$= \frac{\binom{k}{w} \left[\left(\frac{\beta}{\beta+t} \right)^\alpha \right]^k \pi(k)}{C(\alpha, \beta, t, w)}. \quad (2.7)$$

$k \geq w$.

2.4 A estimação de k

Nesta seção discutimos a estimação de k , sob várias distribuições a priori.

Proposição 2.4.1 *Seja π uma distribuição a priori uniforme em $\mathcal{N}^*$, ou seja, $\pi(k) = 1$, $k \in \mathcal{N}^*$. Então,*

(i) *para todo $w \in \mathcal{N}^*$,*

$$\pi(k|\alpha, \beta, t, w) = \binom{k}{w} \left(1 - \left(\frac{\beta}{\beta+t} \right)^\alpha \right)^{w+1} \left(\frac{\beta}{\beta+t} \right)^{\alpha(k-w)}, \quad (2.8)$$

para todo $k \geq w$, ou seja, a distribuição a posteriori de k , dado $W_t = w$, é a distribuição da variável aleatória $w + k'$, onde a variável aleatória k' tem distribuição Binomial Negativa com parâmetros $w + 1$ e $1 - (\beta/(\beta+t))^\alpha$;

(ii) $\pi(k|\alpha, \beta, t, 0) = \pi(k|\alpha, \beta, t, W_t = 0) = (1 - (\frac{\beta}{\beta+t})^\alpha) (\frac{\beta}{\beta+t})^{\alpha(k-1)}$, *para todo $k \geq 1$, isto é, a distribuição a posteriori de k dado $W_t = 0$ é a distribuição da variável aleatória $1 + k''$, onde a variável aleatória k'' tem distribuição Geométrica com parâmetro*

$$1 - \left(\frac{\beta}{\beta+t} \right)^\alpha.$$

Prova Observemos que

$$C(\alpha, \beta, t, w)$$

$$\begin{aligned}
&= \sum_{k \geq w} \binom{k}{w} \left(\frac{\beta}{\beta+t} \right)^{k\alpha} \\
&= \left(\frac{\beta}{\beta+t} \right)^{w\alpha} \sum_{r \geq 0} \binom{-w-1}{r} \left(- \left(\frac{\beta}{\beta+t} \right)^\alpha \right)^r \\
&= \left(\frac{\beta}{\beta+t} \right)^{w\alpha} \left(1 - \left(\frac{\beta}{\beta+t} \right)^\alpha \right)^{-w-1},
\end{aligned} \tag{2.9}$$

o que prova (i). O resultado (ii) segue similarmente. $\diamond$

A média, a variância e a moda da distribuição a posterior de k segue facilmente da proposição 2.4.1

Corolário 2.4.2 *Seja π uma priori uniforme em $\mathcal{N}^*$. Então,*

(i) para todo $w \in \mathcal{N}^$,*

$$E[k|\alpha, \beta, t, w] = E[k|\alpha, \beta, t, W_t = w] = w + \frac{w+1}{\left(\frac{\beta+t}{\beta}\right)^\alpha - 1},$$

$$Var[k|\alpha, \beta, t, w] = Var[k|\alpha, \beta, t, W_t = w]$$

$$= (w+1) \frac{\left(\frac{\beta}{\beta+t}\right)^\alpha}{\left(1 - \left(\frac{\beta}{\beta+t}\right)^\alpha\right)^2};$$

$$(ii) \ E[k|\alpha, \beta, t, 0] = E[k|\alpha, \beta, t, W_t = 0] = 1 + \frac{1}{\left(\frac{\beta+t}{\beta}\right)^\alpha - 1},$$

$$Var[k|\alpha, \beta, t, 0] = Var[k|\alpha, \beta, t, W_t = 0] = \frac{\left(\frac{\beta}{\beta+t}\right)^\alpha}{\left(1 - \left(\frac{\beta}{\beta+t}\right)^\alpha\right)^2};$$

(iii) a moda da distribuição a posteriori de k , dado $W_t = w$, é obtida da expressão (2.8), desde que

$$\frac{\pi(k|\alpha, \beta, t, w)}{\pi(k-1|\alpha, \beta, t, w)} = \begin{cases} > 1, & \text{se } k < \frac{w}{1 - (\frac{\beta}{\beta+t})^\alpha}; \\ = 1, & \text{se } k = \frac{w}{1 - (\frac{\beta}{\beta+t})^\alpha}; \\ < 1, & \text{se } k > \frac{w}{1 - (\frac{\beta}{\beta+t})^\alpha}; \end{cases}$$

para todo $k \geq w + 1$. Assim, se

$$\frac{w}{1 - (\frac{\beta}{\beta+t})^\alpha} \notin \mathcal{N}^*$$

então, a moda a posteriori de k é

$$\lfloor \frac{w}{1 - (\frac{\beta}{\beta+t})^\alpha} \rfloor,$$

onde $\lfloor x \rfloor$ é o maior inteiro menor ou igual a x . Por outro lado, se

$$\frac{w}{1 - (\frac{\beta}{\beta+t})^\alpha} \in \mathcal{N}^*,$$

então as modas da posteriori de k são

$$\frac{w}{1 - (\frac{\beta}{\beta+t})^\alpha} - 1 \quad \text{e} \quad \frac{w}{1 - (\frac{\beta}{\beta+t})^\alpha}.$$

◇

Proposição 2.4.3 *Seja $\{\pi^*(k), k \in \mathcal{N}\}$ uma medida de probabilidade discreta, membro da família exponencial, ou seja,*

$$\pi^*(k) = c(a)S(k)\exp\{d(a)T(k)\},$$

onde $k \in \mathcal{N}$, a é um número real, c e d são funções reais e S, T são funções reais de inteiros positivos. Se π for a função de probabilidade a priori definida por

$$\pi(k) = \frac{\pi^*(k)}{1 - \sum_{i=0}^{k_0} \pi^*(i)},$$

$k \in \{k_0 + 1, k_0 + 2, \dots\}$ onde $k_0 \in \mathcal{N}$, e se $w \in \mathcal{N}$ for tal que $A_{\pi, w} \neq \emptyset$, então a distribuição a posteriori de k dado $W_t = w$ é a distribuição da variável aleatória $w + k'$, onde a distribuição da variável k' , π' , é também membro da família exponencial, ou seja,

$$\pi'(k')$$

$$= c'_1(a, b) c'_2(a, b) S'_1(k') S'_2(k') \exp\{d'_1(a, b) T'_1(k') + d'_2(a, b) T'_2(k')\},$$

$k' \in \{x - w; x \in A_{\pi, w}\}$, onde

$$b = \left(\frac{\beta}{\beta + t}\right)^\alpha;$$

$$c'_1(a, b) = (c(a) \sum_{k \in A_{\pi, w}} \binom{k}{w} b^k S(k) \exp\{d(a) T(k)\})^{-1};$$

$$c'_2(a, b) = c(a); \quad S'_1(k') = S(k' + w); \quad d'_1(a, b) = d(a);$$

$$S'_2(k') = \binom{k' + w}{w}; \quad d'_2(a, b) = \log b; \quad T'_1(k') = T(k' + w)$$

$$\text{e } T'_2(k') = k' + w.$$

Prova. O resultado segue diretamente de (2.7). ◇

Exemplo 2.4.1 Suponha que $\pi^*(k) = p(1 - p)^k$, $k \in \mathcal{N}$, $p \in (0, 1)$, ou seja, π^* é a distribuição de uma variável aleatória com distribuição Geométrica e suponha que $k_0 = 0$. Então, $\pi(k) = p(1 - p)^{k-1}$, $k \in$

$\mathcal{N}^*$, e a distribuição a posteriori de k dado $W_t = w$, $w \in \mathcal{N}^*$, é a distribuição da variável aleatória $w + k'$, onde a distribuição de k' , π' , é

$$\pi'(k') = \binom{k' + w}{w} (1 - (1 - p)(\frac{\beta}{\beta + t})^\alpha)^{w+1} ((1 - p)(\frac{\beta}{\beta + t})^\alpha)^{k'},$$

$k' \in \mathcal{N}$, ou seja, k' tem distribuição Binomial Negativa com parâmetros

$$w + 1 \quad \text{e} \quad 1 - (1 - p)(\frac{\beta}{\beta + t})^\alpha.$$

A distribuição a posteriori de k dado $W_t = 0$ é a distribuição π'' de uma variável aleatória com distribuição Geométrica com parâmetro $1 - (1 - p)(\beta/(\beta + t))^\alpha$, ou seja,

$$\pi''(k') = (1 - (1 - p)(\frac{\beta}{\beta + t})^\alpha) ((1 - p)(\frac{\beta}{\beta + t})^\alpha)^{k'-1},$$

$k' \in \mathcal{N}^*$.

Exemplo 2.4.2 Suponha que

$$\pi^*(k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k \in \mathcal{N}, \quad \lambda > 0,$$

ou seja, π^* é a distribuição de Poisson com parâmetro $\lambda > 0$ e suponha que $k_o = 0$. Então,

$$\pi(k) = \frac{e^{-\lambda} \lambda^k}{k!(1 - e^{-\lambda})}, \quad k \in \mathcal{N}^*,$$

e a distribuição a posteriori de k dado $W_t = w$, $w \in \mathcal{N}^*$, é a distribuição da variável aleatória $w + k'$, onde a distribuição de k' é Poisson com parâmetro

$$\lambda(\frac{\beta}{\beta + t})^\alpha.$$

Além disso, a distribuição a posteriori de k dado $W_t = 0$ é a distribuição de Poisson truncada em zero com parâmetro $\lambda(\frac{\beta}{\beta + t})^\alpha$.

2.5 A estimação de $U(t)$ e $V(t)$

Proposição 2.5.1 *Para quaisquer $a_1, a_2, \dots \in \mathcal{N}$ tal que $w = \sum_{j \geq 1} a_j < \infty$ e para toda priori π tal que $A_{\pi, w} \neq \emptyset$,*

$$(i) \ P[U(t) \leq u | \alpha, \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots] \\ = P[U(t) \leq u | \alpha, \beta, t, W_t = w] = \sum_{k \geq w} G_{(\alpha(k-w); \beta+t)}(u) \pi(k|w),$$

$u > 0$, onde $G_{(0; \beta+t)}(u) = 1$, $u > 0$ e, para $k > w$, $G_{(\alpha(k-w); \beta+t)}(\cdot)$ é a função de distribuição Gama acumulada com parâmetros $\alpha(k-w)$ e $\beta+t$;

$$(ii) \ E[U(t) | \alpha, \beta, t, w] = E[U(t) | \alpha, \beta, t, W_t = w] \\ = \left(\frac{\alpha}{\beta+t} \right) (E[k | \alpha, \beta, t, w] - w) \quad e$$

$$(iii) \ Var[U(t) | \alpha, \beta, t, w] = Var[U(t) | \alpha, \beta, t, W_t = w] \\ = \frac{\alpha}{(\beta+t)^2} \{ E[k | \alpha, \beta, t, w] - w + \alpha Var[k | \alpha, \beta, t, w] \}.$$

Prova.

$$(i) \ P[U(t) \leq u | k, \alpha, \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots] \\ = \sum_{k \geq w} P[U(t) \leq u | \alpha, \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots] \pi(k | \alpha, \beta, t, w) \\ = \sum_{k \geq w} \frac{P[U(t) \leq u, A_1(t) = a_1, A_2(t) = a_2, \dots | k, \alpha, \beta, t]}{P[A_1(t) = a_1, A_2(t) = a_2, \dots | k, \alpha, \beta, t]} \pi(k | \alpha, \beta, t, w) \\ = \pi(w | \alpha, \beta, t, w) \\ + \sum_{k \geq w+1} \frac{P[U(t) \leq u, A_1(t) = a_1, A_2(t) = a_2, \dots | k, \alpha, \beta, t]}{P[A_1(t) = a_1, A_2(t) = a_2, \dots | k, \alpha, \beta, t]} \pi(k | \alpha, \beta, t, w)$$

$$\begin{aligned}
&= \pi(w|\alpha, \beta, t, w) \\
&+ \sum_{k \geq w+1} \left\{ \frac{\sum^* P[U(t) \leq u | k, \alpha, \beta, t, X_1(t) = x_1, \dots, X_k(t) = x_k]}{P[A_1(t) = a_1, A_2(t) = a_2, \dots | k, \alpha, \beta, t]} \right. \\
&\quad \left. \cdot P[X_1(t) = x_1, \dots, X_k(t) = x_k | k, \alpha, \beta, t] \pi(k|\alpha, \beta, t, w) \right\},
\end{aligned}$$

onde $\sum^* = \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]}$. Então, da propriedade (ii) da proposição 2.3.1, segue que

$$\begin{aligned}
&P[U(t) \leq u | \alpha, \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots] = \pi(w|w) \\
&+ \sum_{k \geq w+1} \{G_{(\alpha(k-w); \beta+t)}(u) \cdot \\
&\cdot \frac{\sum^* P[X_1(t) = x_1, \dots, X_k(t) = x_k | \alpha, \beta, t, k] \pi(k|\alpha, \beta, t, w)}{P[A_1(t) = a_1, A_2(t) = a_2, \dots | \alpha, \beta, t, k]} \} \\
&= \pi(w|\alpha, \beta, t, w) + \sum_{k \geq w+1} G_{(\alpha(k-w); \beta+t)}(u) \pi(k|\alpha, \beta, t, w) \\
&= \sum_{k \geq w} G_{(\alpha(k-w); \beta+t)}(u) \pi(k|\alpha, \beta, t, w), \quad u > 0.
\end{aligned}$$

(ii) De (i) segue que

$$\begin{aligned}
E[U(t)|\alpha, \beta, t, w] &= \sum_{k \geq w+1} \int_0^\infty u dG_{(\alpha(k-w); \beta+t)}(u) \pi(k|\alpha, \beta, t, w) \\
&= \sum_{k \geq w} \frac{\alpha(k-w)}{\beta+t} \pi(k|\alpha, \beta, t, w) \\
&= \left(\frac{\alpha}{\beta+t}\right) \sum_{k \geq w} (k-w) \pi(k|\alpha, \beta, t, w) = \left(\frac{\alpha}{\beta+t}\right) (E[k|\alpha, \beta, t, w] - w).
\end{aligned}$$

(iii) De (i) e (ii) segue que

$$\begin{aligned}
Var[U(t)|\alpha, \beta, t, w] &= E[U(t)^2|\alpha, \beta, t, w] - E^2[U(t)|\alpha, \beta, t, w] \\
&= \sum_{k \geq w+1} \int_0^\infty u^2 dG_{(\alpha(k-w); \beta+t)}(u) \pi(k|\alpha, \beta, t, w)
\end{aligned}$$

$$\begin{aligned}
& -\left(\frac{\alpha}{\beta+t}\right)^2(E[k|\alpha, \beta, t, w] - w)^2 \\
& = \sum_{k \geq w} \left\{ \frac{\alpha(k-w)}{(\beta+t)^2} + \left(\frac{\alpha(k-w)}{\beta+t}\right)^2 \right\} \pi(k|\alpha, \beta, t, w) \\
& \quad - \left(\frac{\alpha}{\beta+t}\right)^2(E[k|\alpha, \beta, t, w] - w)^2 \\
& = \frac{\alpha}{(\beta+t)^2} \{E[k|\alpha, \beta, t, w] - w + \alpha \text{Var}[k|\alpha, \beta, t, w]\}.
\end{aligned}$$

◇

Proposição 2.5.2 *Para quaisquer $a_1, a_2, \dots \in \mathcal{N}$ tal que $w = \sum_{j \geq 1} a_j < \infty$ e para toda priori π tal que $A_{\pi, w} \neq \emptyset$,*

$$(i) P[V(t) \leq v | \alpha, \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots]$$

$$= \sum_{k \geq w} \frac{1}{\binom{k}{a_0, a_1, \dots}} \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]}$$

$$P[V(t) \leq v | k, \alpha, \beta, t, ; X_1(t) = x_1, \dots, X_k(t) = x_k] \pi(k|\alpha, \beta, t, w),$$

$$v > 0;$$

$$(ii) E[V(t) | \alpha, \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots]$$

$$= [E[k|\alpha, \beta, t, w] - w] \frac{\alpha}{(\beta+2t)} \left(\frac{\beta+t}{\beta+2t}\right)^\alpha + \sum_{j \geq 1} a_j \frac{(\alpha+j)}{(\beta+2t)} \left(\frac{\beta+t}{\beta+2t}\right)^{\alpha+j}$$

e

$$(iii) \text{Var}[V(t) | \alpha, \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots]$$

$$\begin{aligned}
& = [E[k|\alpha, \beta, t, w] - w] \left\{ \frac{\alpha(\alpha+1)(\beta+t)^\alpha}{(\beta+3t)^{\alpha+2}} - \left[\frac{\alpha(\beta+t)^\alpha}{(\beta+2t)^{\alpha+1}} \right]^2 \right\} \\
& + \sum_{j \geq 1} a_j \left\{ \frac{(\alpha+j)(\alpha+j+1)(\beta+t)^{\alpha+j}}{(\beta+3t)^{\alpha+j+2}} - \left[\frac{(\alpha+j)(\beta+t)^{\alpha+j}}{(\beta+2t)^{\alpha+j+1}} \right]^2 \right\}.
\end{aligned}$$

Prova. Para provar (i) notamos que

$$\begin{aligned}
& P[V(t) \leq v | \alpha, \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots] \\
&= \sum_{k \geq w} \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} P[V(t) \leq v | k, \alpha, \beta, t, X_1(t) = x_1, \dots, X_k(t) = x_k] \\
&\quad \cdot \frac{P[X_1(t) = x_1, \dots, X_k(t) = x_k | \alpha, \beta, t, k]}{P[A_1(t) = a_1, A_2(t) = a_2, \dots | \alpha, \beta, t, k]} \pi(k | \alpha, \beta, t, w).
\end{aligned}$$

O resultado segue de (2.5) e de (vi) na proposição 2.3.1.

Vamos agora provar (ii). Usando (i), podemos escrever

$$\begin{aligned}
& E[V(t) | \alpha, \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots] \tag{2.10} \\
&= \sum_{k \geq w} \frac{1}{\binom{k}{a_0, a_1, \dots}} \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]}
\end{aligned}$$

$$E[V(t) | k, \alpha, \beta, t; X_1(t) = x_1, \dots, X_k(t) = x_k] \pi(k | \alpha, \beta, t, w).$$

De (i) da proposição 2.3.1 segue, para todo $(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]$, que

$$\begin{aligned}
& E[V(t) | k, \alpha, \beta, t; X_1(t) = x_1, \dots, X_k(t) = x_k] \\
&= \sum_{i=1}^k \frac{(\beta + t)^{\alpha + x_i}}{\Gamma(\alpha + x_i)} \int_0^\infty z^{\alpha + x_i} e^{-(\beta + 2t)z} dz \\
&= \sum_{i=1}^k \frac{(\alpha + x_i)}{(\beta + 2t)} \left(\frac{\beta + t}{\beta + 2t} \right)^{\alpha + x_i} \\
&= \sum_{j \geq 0} a_j \frac{(\alpha + j)}{(\beta + 2t)} \left(\frac{\beta + t}{\beta + 2t} \right)^{\alpha + j} \\
&= (k - w) \left(\frac{\alpha}{\beta + 2t} \right) \left(\frac{\beta + t}{\beta + 2t} \right)^\alpha + \sum_{j \geq 1} a_j \left(\frac{\alpha + j}{\beta + 2t} \right) \left(\frac{\beta + t}{\beta + 2t} \right)^{\alpha + j}.
\end{aligned}$$

Assim, (ii) segue de (2.10).

Agora provamos (iii). Para todo $(x_1, \dots, x_k) \in [a_1, a_2, \dots, k]$, segue, usando (i) da proposição 2.3.1, que

$$\begin{aligned}
& Var[V(t)|k, \alpha, \beta, t; X_1(t) = x_1, \dots, X_k(t) = x_k] \\
&= \sum_{i=1}^k \left\{ \int_0^\infty \frac{(\beta+t)^{\alpha+x_i}}{\Gamma(\alpha+x_i)} z^{\alpha+x_i+1} e^{-(\beta+3t)z} dz \right. \\
&\quad \left. - E^2[\lambda_i e^{-\lambda_i t} | k; X_1(t) = x_1, \dots, X_k(t) = x_k] \right\} \\
&= \sum_{i=1}^k \left[\frac{\Gamma(\alpha+x_i+2)}{\Gamma(\alpha+x_i)} \frac{(\beta+t)^{\alpha+x_i}}{(\beta+3t)^{\alpha+x_i+2}} - \left[\frac{(\alpha+x_i)(\beta+t)^{\alpha+x_i}}{(\beta+2t)^{\alpha+x_i+1}} \right]^2 \right] \\
&= \sum_{j=0}^k a_j \left\{ \frac{(\alpha+j)(\alpha+j+1)(\beta+t)^{\alpha+j}}{(\beta+3t)^{\alpha+j+2}} - \left[\frac{(\alpha+j)(\beta+t)^{\alpha+j}}{(\beta+2t)^{\alpha+j+1}} \right]^2 \right\} \\
&= (k-w) \left\{ \frac{\alpha(\alpha+1)(\beta+t)^\alpha}{(\beta+3t)^{\alpha+2}} - \left[\frac{\alpha(\beta+t)^\alpha}{(\beta+2t)^{\alpha+1}} \right]^2 \right\} \\
&\quad + \sum_{j=1}^k a_j \left\{ \frac{(\alpha+j)(\alpha+j+1)(\beta+t)^{\alpha+j}}{(\beta+3t)^{\alpha+j+2}} - \left[\frac{(\alpha+j)(\beta+t)^{\alpha+j}}{(\beta+2t)^{\alpha+j+1}} \right]^2 \right\}.
\end{aligned}$$

Usando (i), temos

$$Var[V(t)|\alpha, \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots]$$

$$= \sum_{k \geq w} \frac{1}{\binom{k}{a_0, a_1, \dots}} \sum_{(x_1, \dots, x_k) \in [a_1 a_2, \dots, k]}$$

$$Var[V(t)|k, \alpha, \beta, t, ; X_1(t) = x_1, \dots, X_k(t) = x_k] \pi(k|\alpha, \beta, t, w),$$

o que conclui a prova. $\diamond$

Exemplo 2.5.1 Usando os resultados do Exemplo 2.4.1, se $\pi(k) = p(1-p)^{k-1}$, $k \in \mathcal{N}^*$, $p \in (0, 1)$ e $w \in \mathcal{N}^*$, então

$$E[k|\alpha, \beta, t, w] = w + \frac{(w+1)(1-p)\left(\frac{\beta}{\beta+t}\right)^\alpha}{1 - (1-p)\left(\frac{\beta}{\beta+t}\right)^\alpha} = \frac{w + (1-p)\left(\frac{\beta}{\beta+t}\right)^\alpha}{1 - (1-p)\left(\frac{\beta}{\beta+t}\right)^\alpha};$$

$$\begin{aligned}
Var[k|\alpha, \beta, t, w] &= \frac{(w+1)(1-p)(\frac{\beta}{\beta+t})^\alpha}{(1 - (1-p)(\frac{\beta}{\beta+t})^\alpha)^2}; \\
E[U(t)|\alpha, \beta, t, w] &= \frac{\alpha}{(\beta+t)} \frac{(w+1)(1-p)(\frac{\beta}{\beta+t})^\alpha}{(1 - (1-p)(\frac{\beta}{\beta+t})^\alpha)}; \\
Var[U(t)|\alpha, \beta, t, w] \\
&= \frac{\alpha}{(\beta+t)^2} \left[\frac{(w+1)(1-p)(\frac{\beta}{\beta+t})^\alpha}{1 - (1-p)(\frac{\beta}{\beta+t})^\alpha} + \alpha \frac{(w+1)(1-p)(\frac{\beta}{\beta+t})^\alpha}{(1 - (1-p)(\frac{\beta}{\beta+t})^\alpha)^2} \right] \\
&= \frac{\alpha(w+1)(1-p)(\frac{\beta}{\beta+t})^\alpha(\alpha + 1 - (1-p)(\frac{\beta}{\beta+t})^\alpha)}{(\beta+t)^2(1 - (1-p)(\frac{\beta}{\beta+t})^\alpha)^2}.
\end{aligned}$$

Se $w = 0$, então

$$\begin{aligned}
E[k|\alpha, \beta, t, 0] &= \frac{1}{1 - (1-p)(\frac{\beta}{\beta+t})^\alpha}; \\
Var[k|\alpha, \beta, t, 0] &= \frac{(1-p)(\frac{\beta}{\beta+t})^\alpha}{(1 - (1-p)(\frac{\beta}{\beta+t})^\alpha)^2}; \\
E[U(t)|\alpha, \beta, t, 0] &= \frac{\alpha}{(\beta+t)(1 - (1-p)(\frac{\beta}{\beta+t})^\alpha)}; \\
Var[U(t)|\alpha, \beta, t, 0] &= \frac{\alpha(1 + (\alpha-1)(1-p)(\frac{\beta}{\beta+t})^\alpha)}{(\beta+t)^2(1 - (1-p)(\frac{\beta}{\beta+t})^\alpha)^2}.
\end{aligned}$$

Além disto,

$$\begin{aligned}
&E[V(t)|\alpha, \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots] \\
&= \frac{(w+1)(1-p)(\frac{\beta}{\beta+t})^\alpha}{1 - (1-p)(\frac{\beta}{\beta+t})^\alpha} \left[\frac{\alpha(\beta+t)^\alpha}{(\beta+2t)^{\alpha+1}} \right] + \sum_{j \geq 1} a_j \frac{(\alpha+j)(\beta+t)^{\alpha+j}}{(\beta+2t)^{\alpha+j+1}},
\end{aligned}$$

e

$$\begin{aligned} & \text{Var}[V(t)|\alpha, \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots] \\ &= \frac{(w+1)(1-p)(\frac{\beta}{\beta+t})^\alpha}{1 - (1-p)(\frac{\beta}{\beta+t})^\alpha} \cdot \left\{ \frac{\alpha(\alpha+1)(\beta+t)^\alpha}{(\beta+3t)^{\alpha+2}} - \left[\frac{\alpha(\beta+t)^\alpha}{(\beta+2t)^{\alpha+1}} \right]^2 \right\} \\ &+ \sum_{j \geq 1} a_j \left\{ \frac{(\alpha+j)(\alpha+j+1)(\beta+t)^{\alpha+j}}{(\beta+3t)^{\alpha+j+2}} - \left[\frac{(\alpha+j)(\beta+t)^{\alpha+j}}{(\beta+2t)^{\alpha+j+1}} \right]^2 \right\}. \end{aligned}$$

Exemplo 2.5.2 De acordo com o Exemplo 2.4.2, se

$$\pi(k) = \frac{e^{-\lambda} \lambda^k}{k!(1 - e^{-\lambda})}, \quad k \in \mathcal{N}^*, \lambda > 0, w \in \mathcal{N}^*,$$

então

$$E[k|\alpha, \beta, t, w] = w + \lambda \left(\frac{\beta}{\beta+t} \right)^\alpha \quad \text{e} \quad \text{Var}[k|\alpha, \beta, t, w] = \lambda \left(\frac{\beta}{\beta+t} \right)^\alpha.$$

Portanto,

$$\begin{aligned} E[U(t)|\alpha, \beta, t, w] &= \left(\frac{\alpha}{\beta+t} \right) \lambda \left(\frac{\beta}{\beta+t} \right)^\alpha \quad \text{e} \\ \text{Var}[U(t)|\alpha, \beta, t, w] &= \frac{\alpha \lambda \beta^\alpha (\alpha+1)}{(\beta+t)^{\alpha+2}}. \end{aligned}$$

Alem disto,

$$\begin{aligned} & E[V(t)|\alpha, \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots] \\ &= \lambda \left(\frac{\beta}{\beta+t} \right)^\alpha \left(\frac{\alpha(\beta+t)^\alpha}{(\beta+2t)^{\alpha+1}} \right) + \sum_{j \geq 1} a_j \frac{(\alpha+j)(\beta+t)^{\alpha+j}}{(\beta+2t)^{\alpha+j+1}} \end{aligned}$$

e

$$\begin{aligned} & \text{Var}[V(t)|\alpha, \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots] \\ &= \lambda \left(\frac{\beta}{\beta+t} \right)^\alpha \cdot \left\{ \frac{\alpha(\alpha+1)(\beta+t)^\alpha}{(\beta+3t)^{\alpha+2}} - \left[\frac{\alpha(\beta+t)^\alpha}{(\beta+2t)^{\alpha+1}} \right]^2 \right\} \\ &+ \sum_{j \geq 1} a_j \left\{ \frac{(\alpha+j)(\alpha+j+1)(\beta+t)^{\alpha+j}}{(\beta+3t)^{\alpha+j+2}} - \left[\frac{(\alpha+j)(\beta+t)^{\alpha+j}}{(\beta+2t)^{\alpha+j+1}} \right]^2 \right\}. \end{aligned}$$

2.6 Resultados assintóticos

Para toda distribuição a priori π para k consideremos a função $\tau : \mathcal{N} \rightarrow \mathcal{N}$, definida por

$$\tau(x) = \begin{cases} \inf A_{\pi,x}, & \text{se } A_{\pi,x} \neq \emptyset, \\ \sup\{n \in \mathcal{N}^*; \pi(n) > 0\}, & \text{se } A_{\pi,x} = \emptyset, \end{cases}$$

para todo $x \in \mathcal{N}$.

Teorema 2.6.1 *Sejam $r \in \mathcal{N}^*$, $a_1, a_2, \dots \in \mathcal{N}$ tal que $w = \sum_{j \geq 1} a_j < \infty$ e π uma distribuição a priori tal que $A_{\pi,w} \neq \emptyset$. Então, quando $t \rightarrow \infty$,*

- (i) $M_r(\alpha, \beta, t, w) = E[k^r | \alpha, \beta, t, W_t = w] \rightarrow [\tau(w)]^r$;
- (ii) $\text{Var}[k | \alpha, \beta, t, W_t = w] \rightarrow 0$;
- (iii) $E[U(t) | \alpha, \beta, t, W_t = w] \rightarrow 0$;
- (iv) $\text{Var}[U(t) | \alpha, \beta, t, W_t = w] \rightarrow 0$;
- (v) $E[V(t) | \alpha, \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots] \rightarrow 0$;
- (vi) $\text{Var}[V(t) | \alpha, \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots] \rightarrow 0$.

Prova. Para simplificar a notação, seja $\tau(w) = \tau$. Então, podemos escrever

$$M_r(\alpha, \beta, t, w) = \frac{\tau^r + A(t)}{1 + B(t)},$$

onde

$$A(t) = \left[\binom{\tau}{w} \left(\frac{\beta}{\beta + t} \right)^{\tau\alpha} \pi(\tau) \right]^{-1} \sum_{n=\tau+1}^{\infty} n^r \binom{n}{w} \left(\frac{\beta}{\beta + t} \right)^{n\alpha} \pi(n)$$

e

$$B(t) = \left[\binom{\tau}{w} \left(\frac{\beta}{\beta + t} \right)^{\tau\alpha} \pi(\tau) \right]^{-1} \sum_{n \geq \tau+1}^{\infty} \binom{n}{w} \left(\frac{\beta}{\beta + t} \right)^{n\alpha} \pi(n).$$

Contudo,

$$\begin{aligned}
\sum_{n=\tau+1}^{\infty} n^r \binom{n}{w} \left(\frac{\beta}{\beta+t}\right)^{n\alpha} \pi(n) &\leq r! \binom{w+r}{w} \sum_{n=\tau+1}^{\infty} \binom{n+r}{w+r} \left(\frac{\beta}{\beta+t}\right)^{n\alpha} \\
&\leq r! \binom{w+r}{w} \frac{(r+\tau+1)!}{(w+r)!} \left(\frac{\beta}{\beta+t}\right)^{\alpha(\tau+1)} \sum_{m=0}^{\infty} \binom{m+r+\tau+1}{r+\tau+1} \left(\frac{\beta}{\beta+t}\right)^{m\alpha} \\
&= \frac{(r+\tau+1)!}{w!} \left(\frac{\beta}{\beta+t}\right)^{\alpha(\tau+1)} \left[1 - \left(\frac{\beta}{\beta+t}\right)^{\alpha}\right]^{-r-\tau-2},
\end{aligned}$$

o que implica que

$$0 \leq A(t) \leq \frac{(r+\tau+1)!}{w!} \frac{\left(\frac{\beta}{\beta+t}\right)^{\alpha}}{\binom{\tau}{w} \pi(\tau)} \left[1 - \left(\frac{\beta}{\beta+t}\right)^{\alpha}\right]^{-r-\tau-2} \rightarrow 0,$$

quando $t \rightarrow \infty$. Como

$$0 \leq B(t) \leq A(t),$$

(i) está provado; segue de (i) que

$$\text{Var}[k|\alpha, \beta, t, w] = E[k^2|\alpha, \beta, t, w] - (E[k|\alpha, \beta, t, w])^2 \rightarrow \tau^2 - \tau^2 = 0,$$

quando $t \rightarrow \infty$, o que prova (ii); (iii) e (iv) seguem diretamente de (i), (ii) e de (ii) e (iii) da Proposição 2.5.1; (v) e (vi) seguem diretamente de (i) e de (ii) e (iii) da Proposição 2.5.2. $\diamond$

Dados $k \in \mathcal{N}^*$ e $\lambda_1, \dots, \lambda_k$ números reais positivos, seja P a medida de probabilidade $P[\cdot|k, \lambda_1, \dots, \lambda_k]$ da superposição dos k processos de Poisson com taxas $\lambda_1, \dots, \lambda_k$.

Lema 2.6.2 *Valem os seguintes resultados:*

(i) $W_t \xrightarrow{q.c.[P]} k$, quando $t \rightarrow \infty$, e

(ii) se π for uma distribuição a priori tal que $A_{\pi,k} \neq \emptyset$, então

$$\tau(W_t) \xrightarrow{q.c.[P]} \tau(k),$$

quando $t \rightarrow \infty$, onde $\xrightarrow{q.c.[P]}$ significa convergência quase certa $[P]$.

Prova. Como para todo $j = 1, 2, \dots, k$,

$$X_j(t) \xrightarrow{q.c.[P]} \infty,$$

quando $t \rightarrow \infty$, então

$$W_t = \sum_{j=1}^k I_{(X_j(t) > 0)} \xrightarrow{q.c.[P]} k,$$

quando $t \rightarrow \infty$, o que prova (i); por outro lado, (ii) segue diretamente de (i). $\diamond$

A prova do próximo teorema segue do Lema 2.6.2.

Teorema 2.6.3 *Seja π uma distribuição a priori tal que $A_{\pi,k} \neq \emptyset$. Então, quando $t \rightarrow \infty$,*

$$(i) \ E[k^r | \alpha, \beta, t, W_t] \xrightarrow{q.c.[P]} [\tau(k)]^r;$$

$$(ii) \ Var[k | \alpha, \beta, t, W_t] \xrightarrow{q.c.[P]} 0;$$

$$(iii) \ E[U(t) | \alpha, \beta, t, W_t] \xrightarrow{q.c.[P]} 0;$$

$$(iv) \ Var[U(t) | \alpha, \beta, t, W_t] \xrightarrow{q.c.[P]} 0;$$

$$(v) \ E[V(t) | \alpha, \beta, t, A_1(t), A_2(t), \dots] \xrightarrow{q.c.[P]} 0;$$

$$(vi) \ Var[V(t) | \alpha, \beta, t, A_1(t), A_2(t), \dots] \xrightarrow{q.c.[P]} 0.$$

Prova. Usando o Teorema 2.6.1, (i), segue que, para todo $r \in \mathcal{N}^*$,

$$E[k^r | \alpha, \beta, t, W_t] = \frac{[\tau(W_t)]^r + A(t, W_t)}{1 + B(t, W_t)},$$

onde

$$0 \leq A(t, W_t) \leq \frac{(\tau(W_t) + r + 1)!}{W_t! \binom{\tau(W_t)}{W_t} \pi(\tau(W_t))} \frac{(\frac{\beta}{\beta+t})^\alpha}{[1 - (\frac{\beta}{\beta+t})^\alpha]^{\tau(W_t)+r+2}}$$

e

$$0 \leq B(t, W_t) \leq A(t, W_t).$$

Como $\tau(W_t) \leq \tau(k)$ q.c. $[P]$, segue que

$$A(t, W_t) \leq \frac{(\tau(k) + r + 1)!}{\min\{\pi(\tau(x)) : 0 \leq x \leq k\}} \frac{(\frac{\beta}{\beta+t})^\alpha}{[1 - (\frac{\beta}{\beta+t})^\alpha]^{\tau(k)+r+2}}, \quad \text{q.c. } [P],$$

o que implica que $A(t, W_t) \rightarrow 0$, q.c. $[P]$, quando $t \rightarrow \infty$. Logo, $B(t, W_t) \rightarrow 0$, q.c. $[P]$, e pelo Lema 2.6.2, $E[k^r | W_t] \rightarrow [\tau(k)]^r$, q.c. $[P]$, quando $t \rightarrow \infty$, provando (i) e (ii); (iii) e (iv) seguem da Proposição 2.5.1 e de (i) e (ii) acima; note que

$$\begin{aligned} 0 &\leq \sum_{j \geq 1} A_j(t) \left(\frac{\alpha + j}{\beta + 2t} \right) \left(\frac{\beta + t}{\beta + 2t} \right)^{\alpha+j} \\ &\leq \frac{(\frac{1}{\beta+2t})(\frac{\beta+t}{\beta+2t})^{\alpha+1} [\alpha(1 - (\frac{\beta+t}{\beta+2t})) + 1] k}{[1 - (\frac{\beta+t}{\beta+2t})]^2} \xrightarrow{\text{q.c. } [P]} 0, \end{aligned}$$

quando $t \rightarrow \infty$. Logo, da Proposição 2.5.2, Lema 2.6.2 e de (i) acima, segue que

$$\begin{aligned} &E[V(t) | \alpha, \beta, t, A_1(t), A_2(t), \dots] \\ &= [E[k | \alpha, \beta, t, W_t] - W_t] \left(\frac{\alpha}{\beta + 2t} \right) \left(\frac{\beta + t}{\beta + 2t} \right)^\alpha \end{aligned}$$

$$+ \sum_{j \geq 1} A_j(t) \left(\frac{\alpha + j}{\beta + 2t} \right) \left(\frac{\beta + t}{\beta + 2t} \right)^{\alpha + j} \xrightarrow{q.c.[P]} 0,$$

quando $t \rightarrow \infty$, o que prova (v): a prova de (vi) segue analogamente.

◊

Capítulo 3

Estimação bayesiana do número de componentes de uma superposição de processos de Poisson: generalizações e propriedades

3.1 Introdução

Consideremos o seguinte procedimento de eliminação das falhas de um novo sistema de software. Inicialmente o sistema tem um número desconhecido, k , de erros $e_1, e_2, \dots, e_k$. O sistema é acionado, exposto a um conjunto variado de problemas, e cada erro e_i , $i = 1, \dots, k$, ocorre independentemente e de acordo com um processo de Poisson com taxa desconhecida λ_i . Assim, os erros vão ocorrendo de acordo com uma superposição de k processos de Poisson homogêneos independentes, com taxas desconhecidas λ_i , $i = 1, \dots, k$. Cada erro gerado e_i é detectado com probabilidade desconhecida q_i , ou não detectado com probabilidade $(1 - q_i)$, $q_i \in (0, 1]$, $i = 1, \dots, k$. Então, se $X_i(t)$ for o número de vezes em que o erro e_i é detectado durante o intervalo de tempo $[0, t]$, temos que, dados k , $\lambda_1, \dots, \lambda_k$ e $q_1, \dots, q_k$,

$\{X_i(t), t \geq 0\}$, $i = 1, 2, \dots, k$, são processos de Poisson homogêneos e independentes com taxas $q_i \lambda_i$, $i = 1, 2, \dots, k$.

O sistema é acionado durante o intervalo de tempo $[0, t]$, $t > 0$, e em seguida, ele é cuidadosamente revisado com a eliminação dos erros detectados. Nosso interesse é estimar

k , número de erros do sistema,

$U(t) = \sum_{i=1}^k \lambda_i I_{(X_i(t)=0)}$, taxa de erro restante do sistema revisado e

$V(t) = E[U(t) | k, q_1, \dots, q_k, \lambda_1, \dots, \lambda_k] = \sum_{i=1}^k \lambda_i e^{-q_i \lambda_i t}$, taxa de erro restante esperada do sistema revisado.

Este problema foi proposto por Ross [57], que estudou a estimação de $U(t)$ e $V(t)$ sob o ponto de vista frequentista. Ele foi também considerado, sob o ponto de vista frequentista, por Nayak [51], [52], [55]. Claramente, temos uma superposição de processos de Poisson como a considerada no capítulo 2, mas, com uma diferença. Lá os erros gerados (detectados com probabilidade 1) eram censurados, no sentido de não identificarem os índices das fontes que os geraram, enquanto que aqui temos uma fonte a mais de censura. Para cada erro gerado durante o intervalo de tempo $[0, t]$ há duas possibilidades: ser detectado ou não ser detectado. Todo erro não detectado é um dado completamente censurado e todo erro detectado é censurado, no sentido de não identificar o índice da fonte que o gerou. Com isto, $X_1(t), X_2(t), \dots, X_k(t)$ não são observáveis. Definamos

$$A_j(t) = \sum_{i=1}^k I_{(X_i(t)=j)}, \text{ para todo } j \in \mathcal{N},$$

ou seja,

$A_j(t)$ é o número de erros detectados exatamente j vezes durante o intervalo de tempo $[0, t]$, com $A_j(0) = 0$, $j = 0, 1, \dots$

As quantidades $A_j(t)$, $j \in \mathcal{N}^*$, são observáveis durante o intervalo de tempo $[0, t]$ e serão usadas para modelar o processo de amostragem de falhas. Observemos que $\sum_{j \geq 0} A_j(t)$ é o número desconhecido de erros do sistema, $W_t = \sum_{j \geq 1} A_j(t)$ é o número de erros distintos observados durante o intervalo de tempo $[0, t]$ e $A_0(t) = k - W_t$ não é observável. Temos então o seguinte problema estatístico: não conhecemos $k, q_1, \dots, q_k, \lambda_1, \dots, \lambda_k$, e queremos estimar $k, U(t)$ e $V(t)$ com base em $A_j(t)$, $j \in \mathcal{N}^*$.

Neste capítulo vamos analisar este problema sob uma formulação bayesiana, supondo que todas as quantidades desconhecidas $k, \lambda_1, \dots, \lambda_k, q_1, \dots, q_k$ são quantidades aleatórias. Evidentemente, esta formulação incorpora a formulação bayesiana considerada no capítulo 2.

Na seção 3.2 apresentamos o modelo de amostragem e a função de verossimilhança. Na seção 3.3 supomos que k tem uma distribuição arbitrária π e que, dado $k, \lambda_1, \dots, \lambda_k, q_1, \dots, q_k$ têm uma distribuição conjunta a priori tal que $q_i \lambda_i$, $i = 1, 2, \dots, k$, são independentes e identicamente distribuídas com função de distribuição F qualquer. Com estas suposições a distribuição a posteriori de k depende dos argumentos F, t e w , onde w é o número de erros distintos detectados durante o intervalo de tempo $[0, t]$. Na seção 3.4 mostramos que a distribuição a posteriori de k é estocasticamente decrescente em F (no sentido de ordenação estocástica) para t e w fixados, estocasticamente crescente em w para F e t fixados e estocasticamente decrescente em t para F e w fixados. Em particular, se F for a função de distribuição Gama com parâmetros α e β , quanto menor for α e maior for β , maior será a distribuição a posteriori de k , para t e w fixados. Na seção 3.5 estudamos a estimação das taxas de falhas restantes do sistema, $U(t)$, e de $V(t)$. Supomos uma priori π qualquer para k e que, dado $k, \lambda_1, \dots, \lambda_k, q_1, \dots, q_k$ são variáveis aleatórias independentes, com $\lambda_1, \dots, \lambda_k$ identicamente distribuídas com distribuição Gama com parâmetros α e

β , e $q_1, \dots, q_k$ identicamente distribuídas com distribuição Beta, com parâmetros α' e α'' , tais que $\alpha' + \alpha'' = \alpha$. As propriedades da distribuição Gama, juntamente com o fato da família de Gamas ser conjugada para amostras Poisson, nos permitirá obter fórmulas explícitas para estimativas de Bayes para $U(t)$ e $V(t)$. Finalmente, na seção 3.6 estudamos o comportamento assintótico dos estimadores dados pelas médias (e variâncias) a posteriori de $U(t)$ e $V(t)$.

3.2 O modelo e a função de verossimilhança

Dados $k, \lambda_1, \dots, \lambda_k, q_1, \dots, q_k$, a distribuição de $A_r(t), r = 1, 2, \dots$ é, para quaisquer $a_1, a_2, \dots \in \mathcal{N}$ tal que $w = \sum_{r \geq 1} a_r \leq k$,

$$P[A_1(t) = a_1, A_2(t) = a_2, \dots | k, \lambda_1, \dots, \lambda_k, q_1, \dots, q_k] \quad (3.1)$$

$$= \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} \prod_{i=1}^k e^{-q_i \lambda_i t} \frac{(q_i \lambda_i t)^{x_i}}{x_i!},$$

onde

$$[a_1, a_2, \dots; k] = \{(y_1, \dots, y_k) \in \mathcal{N}^k : \sum_{i=1}^k I_{\{y_i\}}(r) = a_r, r = 1, 2, \dots\}$$

e

$$\#[a_1, a_2, \dots; k] = \binom{k}{a_0, a_1, \dots},$$

onde $a_0 = k - w$. A função de verossimilhança de $k, \lambda_1, \dots, \lambda_k, q_1, \dots, q_k$, dados que $A_1(t) = a_1, A_2(t) = a_2, \dots$ é dada por

$$L_{(a_1, a_2, \dots)}(k, \lambda_1, \dots, \lambda_k, q_1, \dots, q_k) \quad (3.2)$$

$$= \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} \prod_{i=1}^k e^{-q_i \lambda_i t} \frac{(q_i \lambda_i t)^{x_i}}{x_i!} I_{(0, \infty)}(\lambda_i) I_{(0, 1]}(q_i) I_{\{w, w+1, \dots\}}(k),$$

logo, os dados $a_1, a_2, \dots$ dão alguma informação sobre os parâmetros $q_i \lambda_i$, $i = 1, 2, \dots, k$ e não dão informação sobre $(1 - q_i) \lambda_i$, $i = 1, 2, \dots, k$.

3.3 A distribuição a posteriori de k

Doravante, quando nos referirmos a uma função de distribuição acumulada em $\mathfrak{R}$: conjunto dos números reais, F , fica subentendido que $F(x) = 0$ para todo $x \leq 0$. Seja π uma função de probabilidade em $\mathcal{N}^*$, ou uma distribuição uniforme em $\mathcal{N}^*$, e seja F uma função de distribuição acumulada em $\mathfrak{R}$. Suponhamos que k tenha distribuição a priori π e que, dado k , $\lambda_1, \dots, \lambda_k, q_1, \dots, q_k$ tenham uma distribuição de probabilidade conjunta tal que, dado k , $\lambda'_1, \dots, \lambda'_k$ sejam independentes e identicamente distribuídas de acordo com F , onde $\lambda'_i = q_i \lambda_i$, $i = 1, 2, \dots, k$. Então a função de verossimilhança “marginal” de k , dados $a_1, a_2, \dots$ tais que $\sum_{r \geq 1} a_r < \infty$, é

$$\begin{aligned} & L_{(a_1, a_2, \dots)}(k) \\ &= \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} \prod_{i=1}^k \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^{x_i}}{x_i!} dF(\lambda) I_{\{w, w+1, \dots\}}(k) \\ &= \binom{k}{a_0, a_1, \dots} \prod_{r \geq 0} (q_r(F, t))^{a_r} I_{\{w, w+1, \dots\}}(k), \end{aligned}$$

onde $a_0 = k - w$, $w = \sum_{r \geq 1} a_r$ e

$$q_r(F, t) = \int_0^\infty e^{(-\lambda t)} \frac{(\lambda t)^r}{r!} dF(\lambda), \quad r = 0, 1, \dots$$

Logo,

$$L_{(a_1, a_2, \dots)}(k) \propto \binom{k}{w} (q_0(F, t))^k I_{\{w, w+1, \dots\}}(k) \quad (3.3)$$

onde $q_0(F, t) = \int_0^\infty e^{(-\lambda t)} dF(\lambda)$. Assim, supondo que o sistema seja observado durante um intervalo de tempo $[0, t]$, $t > 0$, e dado que a

amostra é $A_1(t) = a_1, A_2(t) = a_2, \dots$. a distribuição a posteriori de k depende de F, t e w sendo dada por

$$\pi(k|F, t, w) = \frac{\binom{k}{w} (q_0(F, t))^k \pi(k)}{C(F, t, w)}. \quad (3.4)$$

$k \geq w$, onde

$$C(F, t, w) = \sum_{k \geq w} \binom{k}{w} (q_0(F, t))^k \pi(k).$$

Notemos que a distribuição a posteriori de k está bem definida se $A_{\pi, w} \neq \emptyset$, onde para toda priori π e $x \in \mathcal{N}^*$, $A_{\pi, x} = \{n \in \mathcal{N}^* : n \geq x \text{ e } \pi(n) > 0\}$.

Comparando a função de verossimilhança em (3.3) com o *kernel* da função de verossimilhança em (2.6) do capítulo 2, página 23, vemos que, dados π, t e w , a atualização da priori π para a posteriori $\pi(\cdot|F, t, w)$, se dará da mesma forma que em (2.7) do capítulo 2, página 24, independentemente da F . Portanto, valem os seguintes resultados:

- (i) pela Proposição 2.4.1, página 24, se π for uma distribuição a priori uniforme em $\mathcal{N}^*$, ou seja, $\pi(k) = 1, k \in \mathcal{N}^*$, então, para todo $w \in \mathcal{N}^*$,

$$\pi(k|F, t, w) = \binom{k}{w} (1 - q_0(F, t))^{w+1} (q_0(F, t))^{k-w},$$

para todo $k \geq w$, ou seja, a distribuição a posteriori de k , dado $W_t = w$, é a distribuição da variável aleatória $w + k'$, onde a variável aleatória k' tem distribuição Binomial Negativa com parâmetros $w + 1$ e $1 - q_0(F, t)$, e

$$\pi(k|F, t, 0) = (1 - q_0(F, t)) (q_0(F, t))^{k-1},$$

para todo $k \geq 1$, isto é, a distribuição a posteriori de k , dado $W_t = 0$, é a distribuição da variável aleatória $1 + k''$, onde a

variável aleatória k'' tem distribuição Geométrica com parâmetro $1 - q_0(F, t)$:

- (ii) pela Proposição 2.4.3, página 26, se a distribuição a priori para k for membro da família exponencial, então a distribuição a posteriori de k também pertence a família exponencial; mais especificamente, se $\{\pi^*(k), k \in \mathcal{N}\}$ for uma medida de probabilidade discreta, membro da família exponencial, ou seja,

$$\pi^*(k) = c(a)S(k)\exp\{d(a)T(k)\},$$

onde $k \in \mathcal{N}$, a é um número real, c e d são funções reais e S, T são funções reais de inteiros positivos e se $\{\pi(k), k \in \mathcal{N}^*\}$ for a função de probabilidade a priori definida por

$$\pi(k) = \frac{\pi^*(k)}{1 - \sum_{i=0}^{k_0} \pi^*(i)},$$

$k \in \{k_0 + 1, k_0 + 2, \dots\}$ onde $k_0 \in \mathcal{N}$, e se $w \in \mathcal{N}$ for tal que $A_{\pi, w} \neq \emptyset$, então a distribuição a posteriori de k , dado $W_t = w$, é a distribuição da variável aleatória $w + k'$, onde a distribuição da variável k' , π' , é também membro da família exponencial, ou seja,

$$\begin{aligned} & \pi'(k') \\ &= c'_1(a, b)c'_2(a, b)S'_1(k')S'_2(k')\exp\{d'_1(a, b)T'_1(k') + d'_2(a, b)T'_2(k')\}, \\ & k' \in \{x - w; x \in A_{\pi, w}\}, \text{ onde} \end{aligned}$$

$$b = q_0(F, t),$$

$$\begin{aligned} c'_1(a, b) &= (c(a) \sum_{k \in A_{\pi, w}} \binom{k}{w} b^k S(k) \exp\{d(a)T(k)\})^{-1}, \\ c'_2(a, b) &= c(a), \quad S'_1(k') = S(k' + w), \quad d'_1(a, b) = d(a), \end{aligned}$$

$$S'_2(k') = \binom{k' + w}{w}, \quad d'_2(a, b) = \log b, \quad T'_1(k') = T(k' + w)$$

$$\text{e } T'_2(k') = k' + w;$$

- (iii) pelo exemplo 2.4.1, página 27, se $\pi(k) = p(1 - p)^{k-1}$, $k \in \mathcal{N}^*$, então a distribuição a posteriori de k , dado $W_t = w$, é a distribuição da variável aleatória $w + k'$, onde a distribuição de k' é Binomial Negativa com parâmetros

$$w + 1 \quad \text{e} \quad 1 - (1 - p)q_0(F, t);$$

- (iv) pelo exemplo 2.4.2, página 28, se a distribuição a priori para k for Poisson com parâmetro μ , truncada em 1, então a distribuição a posteriori para k , dado $W_t = w$, é igual a distribuição de $w + k'$, onde k' tem distribuição de Poisson truncado do 0 com parâmetro $\mu q_0(F, t)$.

3.4 Propriedades da distribuição a posteriori de k

Nesta seção obtemos importantes propriedades de ordenação da distribuição a posteriori de k e, como consequência imediata, obtemos as correspondentes propriedades de ordenação de estimativas de Bayes de k . Consideramos as estimativas de Bayes dadas pelos momentos a posteriori de k . Para $w \in \mathcal{N}$, $r \in \mathcal{N}^*$, para toda F função de distribuição em $\mathfrak{R}$ e para toda distribuição a priori π para k tal que $A_{\pi, w} \neq \emptyset$, seja $M_r(F, t, w)$ o r -ésimo momento a posteriori de k dado $W_t = w$. Assim,

$$\begin{aligned} M_r(F, t, w) &= E[k^r | F, t, W_t = w] = \sum_{n \geq w} n^r \pi(n | F, t, w) \\ &= \frac{\sum_{n \geq w} n^r \binom{n}{w} (q_0(F, t))^n \pi(n)}{C(F, t, w)}. \end{aligned}$$

Notamos que, para toda distribuição a priori π para k , própria ou imprópria em $\mathcal{N}^*$, $M_r(F, t, w)$ é finito. De fato,

$$\begin{aligned} M_r(F, t, w) &= \frac{\sum_{n=w}^{\infty} n^r \binom{n}{w} (q_0(F, t))^n \pi(n)}{C(F, t, w)} \\ &\leq \frac{\frac{(w+r)!}{w!} \sum_{n=w}^{\infty} \binom{n+r}{w+r} (q_0(F, t))^n}{C(F, t, w)} \\ &= \frac{\frac{(w+r)!}{w!} (q_0(F, t))^w [1 - q_0(F, t)]^{-w-r-1}}{C(F, t, w)} < \infty. \end{aligned}$$

Vamos agora introduzir uma notação. Se F e G são funções de distribuição acumulada das variáveis aleatórias X e Y , respectivamente,

$$F \leq_{estoc} G$$

denota que F é estocasticamente menor do que G , ou seja,

$$P[X > x] \leq P[Y > x], \quad -\infty < x < \infty.$$

Teorema 3.4.1 *Sejam π uma distribuição a priori para k , F e G funções de distribuição acumulada em $\mathfrak{R}$, $t > 0$ e $w \in \mathcal{N}$ tal que $A_{\pi, w} \neq \emptyset$. Então,*

$$F \leq_{estoc} G$$

implica que

$$P[k > x | F, t, w] \geq P[k > x | G, t, w], \quad -\infty < x < \infty.$$

Vale a igualdade se π for degenerada em um ponto $w^ \in A_{\pi, w}$.*

Prova. Seja $x \in \mathfrak{R}$. Dados F e G tais que $F \leq_{estoc} G$, sejam g e h funções reais definidas por

$$g(j) = I_{(x, \infty)}(j) \text{ e}$$

$$h(j) = \left[\frac{q_0(F, t)}{q_0(G, t)} \right]^j,$$

para todo $j \in \mathcal{N}^*$.

Pela caracterização de ordenação estocástica,

$$F \leq_{estoc} G \Leftrightarrow \int_0^\infty f(x) dF(x) \geq \int_0^\infty f(x) dG(x),$$

para toda função não crescente f . Temos então que

$$\int_0^\infty e^{-xt} dF(x) \geq \int_0^\infty e^{-xt} dG(x) \quad \text{ou}$$

$$\frac{q_0(F, t)}{q_0(G, t)} = \frac{\int_0^\infty e^{-xt} dF(x)}{\int_0^\infty e^{-xt} dG(x)} \geq 1$$

o que implica que a função h é não decrescente. Como g é também não decrescente,

$$E[g(k)h(k)|G, t, w] \geq E[g(k)|G, t, w]E[h(k)|G, t, w]. \quad (3.5)$$

Desde que

$$E[g(k)h(k)|G, t, w] = \frac{C(F, t, w)}{C(G, t, w)} E[I_{(k>x)}|F, t, w], \quad (3.6)$$

$$E[g(k)|G, t, w] = E[I_{(k>x)}|G, t, w] \quad (3.7)$$

e

$$E[h(k)|G, t, w] = \frac{C(F, t, w)}{C(G, t, w)}, \quad (3.8)$$

por (3.5), (3.6), (3.7) e (3.8) o resultado está provado. $\diamond$

Corolário 3.4.2 *Sejam π uma distribuição a priori para k , F e G funções de distribuição acumulada em $\mathbb{R}$, $t > 0$, $w \in \mathcal{N}$ tal que $A_{\pi, w} \neq \emptyset$, e $r \in \mathcal{N}^*$. Então,*

$$F \leq_{estoc} G$$

implica que

$$M_r(F, t, w) \geq M_r(G, t, w).$$

Vale a igualdade se π for degenerada em um ponto $w^* \in A_{\pi, w}$.

Prova. Pelo teorema anterior, $F \leq_{estoc} G$ implica que $k|F, t, w$ é estocasticamente maior do que $k|G, t, w$, o que implica que $f(k)|F, t, w$ é estocasticamente maior do que $f(k)|G, t, w$, para toda função crescente f . Logo, basta tomarmos $f(j) = j^r$, $j \in \mathcal{N}^*$. $\diamond$

No capítulo 2, consideramos o caso particular em que F é a função de distribuição Gama com parâmetros α e β . Neste caso, fazendo

$$P[k > x|F, t, w] = P[k > x|\alpha, \beta, t, w], \quad -\infty < x < \infty,$$

e

$$M_r(F, t, w) = M_r(\alpha, \beta, t, w),$$

temos o seguinte corolário.

Corolário 3.4.3 *Sejam π uma distribuição a priori para k , $t > 0$ e $w \in \mathcal{N}$. Então,*

$$\alpha_1 \leq \alpha_2 \quad e \quad \beta_1 \geq \beta_2$$

implicam que

$$P[k > x|\alpha_1, \beta_1, t, w] \geq P[k > x|\alpha_2, \beta_2, t, w], \quad -\infty < x < \infty,$$

e

$$M_r(\alpha_1, \beta_1, t, w) \geq M_r(\alpha_2, \beta_2, t, w).$$

Vale a igualdade se π for degenerada em um ponto $w^* \in A_{\pi, w}$.

Prova. Sejam F_1 e F_2 funções de distribuição acumulada Gama com parâmetros α_1, β_1 e α_2, β_2 , respectivamente. Sabe-se que

$$\alpha_1 \leq \alpha_2 \quad e \quad \beta_1 \geq \beta_2$$

implicam que

$$F_1 \leq_{estoc} F_2,$$

(veja Stoyan e Daley [61, Apêndice 1] ou Taylor [62]). O resultado segue do teorema o corolário anteriores. $\diamond$

Teorema 3.4.4 *Sejam π uma priori para k , F uma função de distribuição acumulada em $\mathbb{R}$, e $w \in \mathcal{N}$ tal que $A_{\pi, w+1} \neq \emptyset$. Então,*

$$P[k > x|F, t, w] \leq P[k > x|F, t, w + 1], \quad -\infty < x < \infty.$$

A igualdade vale se a priori π for degenerada em um ponto $w^* \in A_{\pi, w+1}$.

Prova. Inicialmente notamos que

$$E[I_{(k>x)}(k-w)|F, t, w] - E[I_{(k>x)}|F, t, w]E[k-w|F, t, w] \geq 0, \quad (3.9)$$

desde que $I_{(k>x)}$ e $k-w$ são positivamente correlacionados. Além disto,

$$\begin{aligned} E[I_{(k>x)}(k-w)|F, t, w] &= \frac{\sum_{n=w+1}^{\infty} I_{(x, \infty)}(n)(n-w)\binom{n}{w}(q_0(F, t))^n \pi(n)}{C(F, t, w)} \\ &= (w+1) \frac{C(F, t, w+1)}{C(F, t, w)} E[I_{(k>x)}|F, t, w+1]. \end{aligned} \quad (3.10)$$

Analogamente,

$$E[k-w|w] = (w+1) \frac{C(F, t, w+1)}{C(F, t, w)}. \quad (3.11)$$

De (3.9), (3.10) e (3.11) segue que

$$(w+1) \frac{C(F, t, w+1)}{C(F, t, w)} [E[I_{(k>x)}|F, t, w+1] - E[I_{(k>x)}|F, t, w]] \geq 0,$$

o que implica $E[I_{(k>x)}|F, t, w] \leq E[I_{(k>x)}|F, t, w+1]$. $\diamond$

Corolário 3.4.5 *Sejam π uma priori para k . F uma função de distribuição acumulada em $\mathbb{R}$, $r \in \mathcal{N}^*$ e $w \in \mathcal{N}$ tal que $A_{\pi, w+1} \neq \emptyset$. Então,*

$$M_r(F, t, w) \leq M_r(F, t, w + 1).$$

A igualdade vale se a priori π for degenerada em um ponto $w^ \in A_{\pi, w+1}$.*

Prova. É a mesma do Corolário 3.4.2. ◇

Corolário 3.4.6 *Sejam π uma distribuição a priori qualquer para k , F a função de distribuição acumulada Gama com parâmetros α e β , $r \in \mathcal{N}^*$ e $w \in \mathcal{N}$ tal que $A_{\pi, w+1} \neq \emptyset$. Então,*

$$P[k > x | \alpha, \beta, t, w] \leq P[k > x | \alpha, \beta, t, w + 1], \quad -\infty < x < \infty,$$

e

$$M_r(\alpha, \beta, t, w) \leq M_r(\alpha, \beta, t, w + 1).$$

A igualdade vale se a priori π for degenerada em um ponto $w^ \in A_{\pi, w+1}$.*

Prova. Segue do Teorema 3.4.4 e do Corolário 3.4.5.

Teorema 3.4.7 *Sejam π uma distribuição a priori para k , F uma função de distribuição acumulada em $\mathbb{R}$ e $w \in \mathcal{N}$ tal que $A_{\pi, w} \neq \emptyset$. Então,*

$$s \leq t$$

implica que

$$P[k > x | F, t, w] \leq P[k > x | F, s, w], \quad -\infty < x < \infty,$$

A igualdade vale se π for degenerada em um ponto $w^ \in A_{\pi, w}$.*

Prova. Dados $t \geq s > 0$, $x \in \mathfrak{R}$, sejam g e h funções reais definidas por

$$g(n) = \left[\frac{q_0(F, t)}{q_0(F, s)} \right]^n = \left[\frac{\int_0^\infty e^{-xt} dF(x)}{\int_0^\infty e^{-xs} dF(x)} \right]^n \quad \text{e}$$

$$h(n) = I_{(x, \infty)}(n),$$

para todo $n \in \mathcal{N}^*$. Desde que g é uma função decrescente de n e h é uma função não decrescente de n , segue que

$$E[g(k)|F, s, w]E[h(k)|F, s, w] \geq E[g(k)h(k)|F, s, w]. \quad (3.12)$$

Uma vez que

$$E[g(k)|F, s, w] = \frac{C(F, t, w)}{C(F, s, w)}$$

e

$$E[g(k)h(k)|F, s, w] = \frac{C(F, t, w)}{C(F, s, w)} E[I_{(k > x)}|F, t, w],$$

(3.12) implica que

$$E[I_{(k > x)}|F, t, w] \leq E[I_{(k > x)}|F, s, w].$$

◇

Corolário 3.4.8 *Sejam π uma distribuição a priori para k , F uma função de distribuição acumulada em $\mathfrak{R}$, $r \in \mathcal{N}^*$ e $w \in \mathcal{N}$ tal que $A_{\pi, w} \neq \emptyset$. Então,*

$$s \leq t$$

implica que

$$M_r(F, t, w) \leq M_r(F, s, w).$$

A igualdade vale se π for degenerada em um ponto $w^ \in A_{\pi, w}$.*

Prova. É a mesma do Corolário 3.4.2.

◇

Corolário 3.4.9 *Sejam π uma distribuição a priori para k . F a função de distribuição acumulada Gama com parâmetros α e β , $r \in \mathcal{N}^*$ e $w \in \mathcal{N}$ tal que $A_{\pi,w} \neq \emptyset$. Então,*

$$s \leq t$$

implica que

$$P[k > x | \alpha, \beta, t, w] \leq P[k > x | \alpha, \beta, s, w], \quad -\infty < x < \infty,$$

e

$$M_r(\alpha, \beta, t, w) \leq M_r(\alpha, \beta, s, w).$$

A igualdade vale se π for degenerada em um ponto $w^ \in A_{\pi,w}$.*

Prova. Segue do Teorema 3.4.7 e do Corolário 3.4.8.

Concluindo esta seção observamos que os estimadores dados pelos momentos a posteriori de k são consistentes. Dados $k \in \mathcal{N}^*$, $\lambda_i > 0$ e $q_i \in (0, 1]$, $i = 1, \dots, k$, seja $P = P[\cdot | k, \lambda_1, \dots, \lambda_k, q_1, \dots, q_k]$ a distribuição de $\{(X_1(t), X_2(t), \dots, X_k(t)), t > 0\}$, onde, para todo $t > 0$, $X_i(t)$ é o número de vezes em que o erro e_i é detectado durante o intervalo de tempo $(0, t]$. Pelas nossas suposições anteriores $X_1(t), X_2(t), \dots, X_k(t)$ são independentes e $X_i(t)$ tem distribuição Poisson com parâmetro (ou taxa) $q_i \lambda_i$, $i = 1, 2, \dots, k$. Analogamente ao raciocínio desenvolvido na seção 2.6, página 36, temos que $W(t) \rightarrow k$ q.c. $[P]$, $\tau(W(t)) \rightarrow \tau(k)$ q.c. $[P]$, quando $t \rightarrow \infty$, onde

$$\tau(k) = \begin{cases} \inf A_{\pi,k}, & \text{se } A_{\pi,k} \neq \emptyset, \\ \sup\{n \in \mathcal{N}^*; \pi(n) > 0\}, & \text{se } A_{\pi,k} = \emptyset, \end{cases}$$

e π é uma distribuição a priori qualquer para k . Por outro lado, podemos demonstrar, de modo análogo a demonstração feita no item (i) do Teorema 2.6.3, página 38, que

$$E[k^r | F, t, W(t)] \rightarrow [\tau(k)]^r, \quad \text{q.c.}[P].$$

3.5 A estimação de $U(t)$ e $V(t)$

Nesta seção vamos reparametrizar nosso modelo. Esta reparametrização nos proporcionará uma melhor visualização da estimação de $U(t)$ e $V(t)$.

Sejam

$\lambda'_i = q_i \lambda_i$, taxa de ocorrência dos erros e_i detectados e

$\lambda''_i = (1 - q_i) \lambda_i$, taxa de ocorrência dos erros e_i não detectados, $i = 1, 2, \dots, k$.

Sob esta reparametrização, dados $k, \lambda'_1, \dots, \lambda'_k$ e $\lambda''_1, \dots, \lambda''_k$,

$\{X_i(t), t \geq 0\}$, $i = 1, 2, \dots, k$, são processos de Poisson homogêneos e independentes com taxas λ'_i , $i = 1, 2, \dots, k$,

e as quantidades de interesse $U(t)$ e $V(t)$ podem ser reescritas como,

$$U(t) = \sum_{i=1}^k \lambda'_i I_{\{X_i(t)=0\}} + \sum_{i=1}^k \lambda''_i I_{\{X_i(t)=0\}}$$

e

$$V(t) = \sum_{i=1}^k \lambda'_i e^{-\lambda'_i t} + \sum_{i=1}^k \lambda''_i e^{-\lambda''_i t}.$$

Podemos ver, pelas igualdades acima, como $U(t)$ e $V(t)$ dependem das taxas dos processos relativos aos erros detectados e das taxas dos processos relativos aos erros não detectados. Dados $k, \lambda'_1, \dots, \lambda'_k, \lambda''_1, \dots, \lambda''_k$, a distribuição de $A_1(t), A_2(t), \dots$ é

$$\begin{aligned} P[A_1(t) = a_1, A_2(t) = a_2, \dots | k, \lambda'_1, \dots, \lambda'_k, \lambda''_1, \dots, \lambda''_k] & \quad (3.13) \\ = \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots, k]} \prod_{i=1}^k e^{-\lambda'_i t} \frac{(\lambda'_i t)^{x_i}}{x_i!}, \end{aligned}$$

para quaisquer $a_1, a_2, \dots \in \mathcal{N}$, tal que $w = \sum_{r \geq 1} a_r \leq k$. Logo, a função de verossimilhança de $k, \lambda'_1, \dots, \lambda'_k, \lambda''_1, \dots, \lambda''_k$, dados que $A_1(t) = a_1, A_2(t) = a_2, \dots$, é

$$L_{(a_1, a_2, \dots)}(k, \lambda'_1, \dots, \lambda'_k, \lambda''_1, \dots, \lambda''_k) \quad (3.14)$$

$$= \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} \prod_{i=1}^k e^{-\lambda'_i t} \frac{(\lambda'_i t)^{x_i}}{x_i!} I_{(0, \infty)}(\lambda'_i) I_{(0, \infty)}(\lambda''_i) I_{\{w, w+1, \dots\}}(k),$$

nos permite concluir que os dados $a_1, a_2, \dots$ dão alguma informação sobre as taxas dos processos relativos aos erros detectáveis λ'_i , $i = 1, 2, \dots, k$, mas não dão informação alguma sobre as taxas dos processos relativos aos erros não detectáveis λ''_i , $i = 1, 2, \dots, k$.

Seja π uma função de probabilidade qualquer ou uma distribuição imprópria uniforme em $\mathcal{N}^*$. Nesta seção, supomos que k tem distribuição π e que, dado k , $k \in \mathcal{N}^*$, $\lambda_1, \dots, \lambda_k, q_1, \dots, q_k$ são variáveis aleatórias independentes, com $\lambda_1, \dots, \lambda_k$ identicamente distribuídas com distribuição Gama com parâmetros α e β , $\alpha > 0$, $\beta > 0$, e $q_1, \dots, q_k$ identicamente distribuídas com distribuição Beta com parâmetros α' e α'' , $\alpha' > 0$, $\alpha'' > 0$ tais que $\alpha' + \alpha'' = \alpha$. Então, a distribuição a priori conjunta de $(\lambda_1, \dots, \lambda_k, q_1, \dots, q_k, k)$ é

$$\begin{aligned} \pi(\lambda_1, \dots, \lambda_k, q_1, \dots, q_k, k) &= \pi(\lambda_1, \dots, \lambda_k, q_1, \dots, q_k | k) \pi(k) \\ &= \prod_{i=1}^k \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda_i^{\alpha-1} e^{-\beta \lambda_i} I_{(0, \infty)}(\lambda_i) \frac{q_i^{\alpha'-1} (1 - q_i)^{\alpha''-1}}{B(\alpha', \alpha'')} I_{(0, 1]}(q_i) \pi(k) I_{\mathcal{N}^*}(k). \end{aligned} \quad (3.15)$$

Devido as propriedades da distribuição Gama, esta suposição é equivalente a suposição de que k tem distribuição π e que, dado k , $\lambda'_1, \dots, \lambda'_k, \lambda''_1, \dots, \lambda''_k$ são variáveis aleatórias independentes, com $\lambda'_1, \dots, \lambda'_k$ identicamente distribuídas com distribuição Gama com parâmetros α' e β e $\lambda''_1, \dots, \lambda''_k$ identicamente distribuídas com distribuição Gama com

parâmetros α'' e β . Logo, a distribuição a priori (3.15) pode ser equivalentemente especificada pela distribuição a priori

$$\begin{aligned} \pi(\lambda'_1, \dots, \lambda'_k, \lambda''_1, \dots, \lambda''_k, k) &= \pi(\lambda'_1, \dots, \lambda'_k, \lambda''_1, \dots, \lambda''_k | k) \pi(k) \\ &= \prod_{i=1}^k \frac{\beta^{\alpha'}}{\Gamma(\alpha')} \lambda_i'^{\alpha'-1} e^{-\beta \lambda_i'} I_{(0,\infty)}(\lambda_i') \frac{\beta^{\alpha''}}{\Gamma(\alpha'')} \lambda_i''^{\alpha''-1} e^{-\beta \lambda_i''} I_{(0,\infty)}(\lambda_i'') \pi(k) I_{\mathcal{N}^*}(k). \end{aligned} \quad (3.16)$$

Lema 3.5.1

(i) dados k e $X_1(t) = x_1, \dots, X_k(t) = x_k$, $\lambda'_1, \dots, \lambda'_k, \lambda''_1, \dots, \lambda''_k$ são variáveis aleatórias independentes, onde

λ'_i tem distribuição Gama com parâmetros $\alpha' + x_i$ e $\beta + t$ e

λ''_i tem distribuição Gama com parâmetros α'' e β , $i = 1, 2, \dots, k$;

(ii) dados k e $X_1(t) = x_1, \dots, X_k(t) = x_k$ tal que $w = \sum_{i=1}^k I_{\mathcal{N}^*}(x_i) < k$,

$$U(t) = \sum_{i=1}^k \lambda'_i I_{(X_i(t)=0)} + \sum_{i=1}^k \lambda''_i I_{(X_i(t)=0)},$$

onde

$\sum_{i=1}^k \lambda'_i I_{(X_i(t)=0)}$ e $\sum_{i=1}^k \lambda''_i I_{(X_i(t)=0)}$ são independentes,

$\sum_{i=1}^k \lambda'_i I_{(X_i(t)=0)}$ tem distribuição Gama com parâmetros $(k-w)\alpha'$ e $(\beta + t)$, e

$\sum_{i=1}^k \lambda''_i I_{(X_i(t)=0)}$ tem distribuição Gama com parâmetros $(k-w)\alpha''$ e β . Portanto,

$$\begin{aligned} E[U(t) | k, \alpha', \alpha'', \beta, t, x_1, \dots, x_k] \\ = (k-w) \frac{\alpha'}{\beta + t} + (k-w) \frac{\alpha''}{\beta} \end{aligned}$$

e

$$\begin{aligned} & Var[U(t)|k, \alpha', \alpha'', \beta, t, x_1, \dots, x_k] \\ &= (k-w) \frac{\alpha'}{(\beta+t)^2} + (k-w) \frac{\alpha''}{\beta^2}; \end{aligned}$$

(iii) dados k e $X_1(t) = x_1, \dots, X_k(t) = x_k$ tal que $w = \sum_{i=1}^k I_{N^*}(x_i) < k$,

$$\begin{aligned} & E[V(t)|k, \alpha', \alpha'', \beta, t, x_1, \dots, x_k] \\ &= (k-w) \frac{\alpha'}{(\beta+2t)} \left(\frac{\beta+t}{\beta+2t} \right)^{\alpha'} + \sum_{j \geq 1} a_j \left(\frac{\alpha'+j}{\beta+2t} \right) \left(\frac{\beta+t}{\beta+2t} \right)^{\alpha'+j} + \\ &+ (k-w) \frac{\alpha''}{\beta} \left(\frac{\beta+t}{\beta+2t} \right)^{\alpha'} + \sum_{j \geq 1} a_j \left(\frac{\alpha''}{\beta} \right) \left(\frac{\beta+t}{\beta+2t} \right)^{\alpha'+j} \end{aligned}$$

e

$$\begin{aligned} & Var[V(t)|k, \alpha', \alpha'', \beta, t, x_1, \dots, x_k] \\ &= (k-w) \left\{ \left(\frac{\alpha'+1}{\beta+3t} \right) \left(\frac{\alpha'}{\beta+3t} \right) \left(\frac{\beta+t}{\beta+3t} \right)^{\alpha'} - \left(\frac{\alpha'}{\beta+2t} \right)^2 \left(\frac{\beta+t}{\beta+2t} \right)^{2\alpha'} + \right. \\ &+ \left(\frac{\alpha''}{\beta} \right)^2 \left[\left(\frac{\beta+t}{\beta+3t} \right)^{\alpha'} - \left(\frac{\beta+t}{\beta+2t} \right)^{2\alpha'} \right] + \frac{\alpha''}{\beta^2} \left(\frac{\beta+t}{\beta+3t} \right)^{\alpha'} + \\ &+ 2 \left(\frac{\alpha''}{\beta} \right) \left[\left(\frac{\alpha'+1}{\beta+3t} \right) \left(\frac{\alpha'}{\beta+3t} \right) \left(\frac{\beta+t}{\beta+3t} \right)^{\alpha'} - \left(\frac{\alpha'}{\beta+2t} \right) \left(\frac{\beta+t}{\beta+2t} \right)^{\alpha'} \left(\frac{\beta+t}{\beta+2t} \right)^{\alpha'} \right] \Big\} \\ &+ \sum_{j \geq 1} a_j \left\{ \left(\frac{\alpha'+j+1}{\beta+3t} \right) \left(\frac{\alpha'+j}{\beta+3t} \right) \left(\frac{\beta+t}{\beta+3t} \right)^{\alpha'+j} - \left(\frac{\alpha'+j}{\beta+2t} \right)^2 \left(\frac{\beta+t}{\beta+2t} \right)^{2(\alpha'+j)} + \right. \end{aligned}$$

$$\begin{aligned}
& + \left(\frac{\alpha''}{\beta}\right)^2 \left[\left(\frac{\beta+t}{\beta+3t}\right)^{\alpha'+j} - \left(\frac{\beta+t}{\beta+2t}\right)^{2\alpha'+j} \right] + \frac{\alpha''}{\beta^2} \left(\frac{\beta+t}{\beta+3t}\right)^{\alpha'+j} + \\
& + 2\left(\frac{\alpha''}{\beta}\right) \left[\left(\frac{\alpha'+j}{\beta+3t}\right) \left(\frac{\beta+t}{\beta+3t}\right)^{\alpha'+j} - \left(\frac{\alpha'+j}{\beta+2t}\right) \left(\frac{\beta+t}{\beta+2t}\right)^{\alpha'+j} \left(\frac{\beta+t}{\beta+2t}\right)^{\alpha'+j} \right],
\end{aligned}$$

onde $(k-w) = \sum_{i=1}^k I_{\{0\}}(x_i)$ e $a_j = \sum_{i=1}^k I_{\{j\}}(x_i)$, $j = 1, 2, \dots$;

(iv) dado k e dado que $A_1(t) = a_1, A_2(t) = a_2, \dots$, a distribuição de $(X_1(t), \dots, X_k(t))$ é uniforme no conjunto $[a_1, a_2, \dots; k]$, ou seja,

$$\begin{aligned}
& P[X_1(t) = x_1, \dots, X_k(t) = x_k | k; A_1(t) = a_1, A_2(t) = a_2, \dots] \\
& = \frac{1}{\binom{k}{w} \binom{w}{a_1, a_2, \dots}} I_{[a_1, a_2, \dots; k]}(x_1, \dots, x_k).
\end{aligned}$$

Prova. Para provarmos (i) observemos que

$$\begin{aligned}
& \pi(\lambda'_1, \dots, \lambda'_k, \lambda''_1, \dots, \lambda''_k | k, x_1, \dots, x_k) \\
& \propto P[x_1, \dots, x_k | k, \lambda'_1, \dots, \lambda'_k, \lambda''_1, \dots, \lambda''_k] \pi(\lambda'_1, \dots, \lambda'_k, \lambda''_1, \dots, \lambda''_k | k) \\
& = P[x_1, \dots, x_k | k, \lambda'_1, \dots, \lambda'_k] \pi(\lambda'_1, \dots, \lambda'_k | k) \pi(\lambda''_1, \dots, \lambda''_k | k) \\
& \propto \pi(\lambda'_1, \dots, \lambda'_k | k, x_1, \dots, x_k) \pi(\lambda''_1, \dots, \lambda''_k | k).
\end{aligned}$$

Mas, pela Proposição 2.3.1 (i), dado k e dado que $X_1(t) = x_1, \dots, X_k(t) = x_k$, $\lambda'_1, \dots, \lambda'_k$ são independentes com λ'_i tendo distribuição Gama com parâmetros $\alpha' + x_i$ e $\beta + t$, $i = 1, 2, \dots, k$ e, por hipótese, dado k , $\lambda''_1, \dots, \lambda''_k$ são variáveis aleatórias independentes identicamente distribuídas com distribuição Gama com parâmetros α'' e β , o que prova (i). O item (ii) segue de (i) e da propriedade reprodutiva da distribuição Gama. A seguir provamos (iii).

$$E[V(t) | k, \alpha', \alpha'', \beta, t, x_1, \dots, x_k]$$

$$\begin{aligned}
&= E\left[\sum_{i=1}^k (\lambda'_i + \lambda''_i) e^{-\lambda'_i t} | k, \alpha', \alpha'', \beta, t, x_1, \dots, x_k\right] \\
&= \sum_{i=1}^k E[(\lambda'_i + \lambda''_i) e^{-\lambda'_i t} | k, \alpha', \alpha'', \beta, t, x_1, \dots, x_k].
\end{aligned}$$

Pelo item (i) acima, dados k e $x_1, \dots, x_k$, λ'_i tem distribuição Gama com parâmetros $\alpha' + x_i$ e $\beta + t$, é independente de λ''_i , que tem distribuição Gama com parâmetros α'' e β , $i = 1, 2, \dots, k$. Então,

$$\begin{aligned}
&E[V(t) | k, \alpha', \alpha'', \beta, t, x_1, \dots, x_k] \\
&= \sum_{i=1}^k E[\lambda'_i e^{-\lambda'_i t} | k, \alpha', \alpha'', \beta, t, x_1, \dots, x_k] + E[\lambda''_i] E[e^{-\lambda'_i t} | k, \alpha', \alpha'', \beta, t, x_1, \dots, x_k] \\
&= \sum_{i=1}^k \frac{(\alpha' + x_i)}{(\beta + 2t)} \left(\frac{\beta + t}{\beta + 2t}\right)^{\alpha' + x_i} + \frac{\alpha''}{\beta} \sum_{i=1}^k \left(\frac{\beta + t}{\beta + 2t}\right)^{\alpha' + x_i}.
\end{aligned}$$

Do item (i) acima temos que dados k e $x_1, \dots, x_k$, $(\lambda'_i, \lambda''_i)$, $i = 1, 2, \dots, k$, são independentes. Então,

$$\begin{aligned}
&Var[V(t) | k, \alpha', \alpha'', \beta, t, x_1, \dots, x_k] \\
&= Var\left[\sum_{i=1}^k (\lambda'_i + \lambda''_i) e^{-\lambda'_i t} | k, \alpha', \alpha'', \beta, t, x_1, \dots, x_k\right] \\
&= \sum_{i=1}^k Var[(\lambda'_i + \lambda''_i) e^{-\lambda'_i t} | k, \alpha', \alpha'', \beta, t, x_1, \dots, x_k] \\
&= \sum_{i=1}^k \{Var[\lambda'_i e^{-\lambda'_i t} | k, \alpha', \alpha'', \beta, t, x_1, \dots, x_k] + Var[\lambda''_i e^{-\lambda'_i t} | k, \alpha', \alpha'', \beta, t, x_1, \dots, x_k] + \\
&\quad + 2Cov[\lambda'_i e^{-\lambda'_i t}, \lambda''_i e^{-\lambda'_i t} | k, \alpha', \alpha'', \beta, t, x_1, \dots, x_k]\}.
\end{aligned}$$

De (i) acima, dados k e $x_1, \dots, x_k$, $\lambda'_1, \dots, \lambda'_k, \lambda''_1, \dots, \lambda''_k$ são independentes, λ'_i tendo distribuição Gama com parâmetros $\alpha' + x_i$ e $\beta + t$, λ''_i tendo distribuição Gama com parâmetros α'' e β . Então,

$$Var[V(t) | k, \alpha', \alpha'', \beta, t, x_1, \dots, x_k]$$

$$\begin{aligned}
&= \sum_{i=1}^k \left\{ \left(\frac{\alpha' + x_i + 1}{\beta + 3t} \right) \left(\frac{\alpha' + x_i}{\beta + 3t} \right) \left(\frac{\beta + t}{\beta + 3t} \right)^{\alpha' + x_i} - \left(\frac{\alpha' + x_i}{\beta + 2t} \right)^2 \left(\frac{\beta + t}{\beta + 2t} \right)^{2(\alpha' + x_i)} + \right. \\
&\quad \left. + \left(\frac{\alpha''}{\beta} \right)^2 \left[\left(\frac{\beta + t}{\beta + 3t} \right)^{\alpha' + x_i} - \left(\frac{\beta + t}{\beta + 2t} \right)^{2(\alpha' + x_i)} \right] + \frac{\alpha''}{\beta^2} \left(\frac{\beta + t}{\beta + 3t} \right)^{\alpha' + x_i} + \right. \\
&\quad \left. + 2 \left(\frac{\alpha''}{\beta} \right) \left[\left(\frac{\alpha' + x_i}{\beta + 3t} \right) \left(\frac{\beta + t}{\beta + 3t} \right)^{\alpha' + x_i} - \left(\frac{\alpha' + x_i}{\beta + 2t} \right) \left(\frac{\beta + t}{\beta + 2t} \right)^{\alpha' + x_i} \left(\frac{\beta + t}{\beta + 2t} \right)^{\alpha' + x_i} \right] \right\},
\end{aligned}$$

o que prova (iii). Para provar (iv) basta observarmos que, dado k , $X_1(t), \dots, X_k(t)$ são independentes e identicamente distribuídas com distribuição Binomial Negativa com parâmetros α' e $\beta/(\beta + t)$, ou seja,

$$P[X_i(t) = x|k] = \frac{\Gamma(\alpha' + x)}{\Gamma(\alpha')x!} \left(\frac{\beta}{\beta + t} \right)^{\alpha'} \left(\frac{t}{\beta + t} \right)^x I_{\mathcal{N}}(x),$$

$i = 1, 2, \dots, k$, e portanto,

$$\begin{aligned}
&P[X_1(t) = x_1, \dots, X_k(t) = x_k|k] \\
&= \left(\frac{\beta}{\beta + t} \right)^{k\alpha'} \prod_{i=1}^k \frac{\Gamma(\alpha' + x_i)}{\Gamma(\alpha')x_i!} \left(\frac{t}{\beta + t} \right)^{x_i} I_{\mathcal{N}}(x_i) \\
&= \left(\frac{\beta}{\beta + t} \right)^{k\alpha'} \prod_{r \geq 1} \left\{ \frac{\Gamma(\alpha' + r)}{\Gamma(\alpha')r!} \left(\frac{t}{\beta + t} \right)^r I_{\mathcal{N}}(r) \right\}^{a_r},
\end{aligned}$$

para todo $(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]$.

◇

Proposição 3.5.2 *Para quaisquer $a_1, a_2, \dots \in \mathcal{N}$, tal que $w = \sum_{j \geq 1} a_j < \infty$ e para toda priori π tal que $A_{\pi, w} \neq \emptyset$,*

$$\begin{aligned}
(i) \quad &P[U(t) \leq u | \alpha', \alpha'', \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots] \\
&= P[U(t) \leq u | \alpha', \alpha'', \beta, t, W_t = w]
\end{aligned}$$

$$= \sum_{k \geq w} G_{((k-w)\alpha'; \beta+t)} * G_{((k-w)\alpha''; \beta)}(u) \pi(k|\alpha', \beta, t, w),$$

$u > 0$. onde $G_{((k-w)\alpha'; \beta+t)} * G_{((k-w)\alpha''; \beta)}$ é a convolução de duas funções de distribuição Gama independentes, onde $G_{(0, \beta+t)}(u) = G_{(0, \beta)}(u) = 0$, $u > 0$, para $k > w$, $G_{((k-w)\alpha'; \beta+t)}(\cdot)$ é a função de distribuição Gama com parâmetros $(k-w)\alpha'$ e $\beta+t$, e $G_{((k-w)\alpha''; \beta)}(\cdot)$ é a função de distribuição Gama com parâmetros $(k-w)\alpha''$ e β ;

(ii)

$$\begin{aligned} & E[U(t)|\alpha', \alpha'', \beta, t, w] \\ &= (E[k|\alpha', \beta, t, w] - w) \frac{\alpha'}{\beta + t} + (E[k|\alpha'', \beta, t, w] - w) \frac{\alpha''}{\beta}; \end{aligned}$$

(iii)

$$\begin{aligned} & Var[U(t)|\alpha', \alpha'', \beta, t, w] \\ &= (E[k|\alpha', \beta, t, w] - w) \frac{\alpha'}{(\beta + t)^2} + (E[k|\alpha'', \beta, t, w] - w) \frac{\alpha''}{\beta^2} \\ & \quad + Var[k|\alpha', \alpha'', \beta, t, w] \{(\alpha'/\beta + t)^2 + (\alpha''/\beta)^2\}. \end{aligned}$$

Prova. Seja $u > 0$.

$$P[U(t) \leq u|\alpha', \alpha'', \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots]$$

$$= \sum_{k \geq w} P[U(t) \leq u|k; \alpha', \alpha'', \beta, A_1(t) = a_1, A_2(t) = a_2, \dots].$$

$$\cdot \pi(k|\alpha', \alpha'', \beta, A_1(t) = a_1, A_2(t) = a_2, \dots)$$

$$= \pi(w|\alpha', \beta, t, w) + \sum_{k \geq w+1}^* \{ \sum$$

$$P[U(t) \leq u|k; \alpha', \alpha'', \beta, t, X_1(t) = x_1, \dots, X_k(t) = x_k, A_1(t) = a_1, \dots].$$

$$\cdot P[X_1(t) = x_1, \dots, X_k(t) = x_k|k; \alpha', \alpha'', \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots].$$

$$\begin{aligned}
& \cdot \pi(k|\alpha', \beta, t, w)\} \\
& = \pi(w|\alpha', \alpha'', \beta, t, w) + \\
& + \sum_{k \geq w+1}^* \{ \sum P[U(t) \leq u|k; \alpha', \alpha'', \beta, t, X_1(t) = x_1, \dots, X_k(t) = x_k].
\end{aligned}$$

$$\begin{aligned}
& \cdot P[X_1(t) = x_1, \dots, X_k(t) = x_k|k; \alpha', \alpha'', \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots] \\
& \pi(k|\alpha', \beta, t, w)\},
\end{aligned}$$

onde $\Sigma^* = \Sigma_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]}$. Pelo Lema 3.5.1 (ii) e Lema 3.5.1 (iv) provamos (i). Os itens (ii) e (iii) seguem diretamente do item (i) acima e do Lema 3.5.1 (ii).

Proposição 3.5.3 *Para quaisquer $a_1, a_2, \dots \in \mathcal{N}$ tal que $w = \sum_{j \geq 1} a_j < \infty$ e para toda priori π tal que $A_{\pi, w} \neq \emptyset$,*

(i)

$$\begin{aligned}
& E[V(t)|\alpha', \alpha'', \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots] \\
& = (E[k|\alpha', \beta, t, w] - w) \frac{\alpha'}{(\beta + 2t)} \left(\frac{\beta + t}{\beta + 2t} \right)^{\alpha'} + \sum_{j \geq 1} a_j \left(\frac{\alpha' + j}{\beta + 2t} \right) \left(\frac{\beta + t}{\beta + 2t} \right)^{\alpha' + j} + \\
& + (E[k|\alpha', \beta, t, w] - w) \frac{\alpha''}{\beta} \left(\frac{\beta + t}{\beta + 2t} \right)^{\alpha'} + \sum_{j \geq 1} a_j \left(\frac{\alpha''}{\beta} \right) \left(\frac{\beta + t}{\beta + 2t} \right)^{\alpha' + j} \\
& e
\end{aligned}$$

(ii)

$$\begin{aligned}
& Var[V(t)|\alpha', \alpha'', \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots] \\
& = (E[k|\alpha', \beta, t, w] - w) \left\{ \frac{(\alpha' + 1)}{(\beta + 3t)} \frac{\alpha'}{(\beta + 3t)} \left(\frac{\beta + t}{\beta + 3t} \right)^{\alpha'} - \left(\frac{\alpha'}{(\beta + 2t)} \left(\frac{\beta + t}{\beta + 2t} \right)^{\alpha'} \right)^2 \right\} +
\end{aligned}$$

$$\begin{aligned}
& + \left(\frac{\alpha''}{\beta}\right)^2 \left[\left(\frac{\beta+t}{\beta+3t}\right)^{\alpha'} - \left(\frac{\beta+t}{\beta+2t}\right)^{2\alpha'} \right] + \frac{\alpha'}{\beta^2} \left(\frac{\beta+t}{\beta+3t}\right)^{\alpha'} + \\
& + 2\left(\frac{\alpha''}{\beta}\right) \left[\frac{(\alpha'+1)}{(\beta+3t)} \frac{\alpha'}{(\beta+3t)} \left(\frac{\beta+t}{\beta+3t}\right)^{\alpha'} - \frac{\alpha'}{(\beta+2t)} \left(\frac{\beta+t}{\beta+2t}\right)^{\alpha'} \left(\frac{\beta+t}{\beta+2t}\right)^{\alpha'} \right] \} \\
& + \sum_{j \geq 1} a_j \left\{ \frac{(\alpha'+j+1)}{(\beta+3t)} \frac{\alpha'+j}{(\beta+3t)} \left(\frac{\beta+t}{\beta+3t}\right)^{\alpha'+j} - \left(\frac{\alpha'+j}{(\beta+2t)} \left(\frac{\beta+t}{\beta+2t}\right)^{\alpha'+j}\right)^2 + \right. \\
& \quad \left. + \left(\frac{\alpha''}{\beta}\right)^2 \left[\left(\frac{\beta+t}{\beta+3t}\right)^{\alpha'+j} - \left(\frac{\beta+t}{\beta+2t}\right)^{2\alpha'+j} \right] + \frac{\alpha'}{\beta^2} \left(\frac{\beta+t}{\beta+3t}\right)^{\alpha'+j} + \right. \\
& \quad \left. + 2\left(\frac{\alpha''}{\beta}\right) \left[\frac{(\alpha'+j+1)}{(\beta+3t)} \frac{\alpha'+j}{(\beta+3t)} \left(\frac{\beta+t}{\beta+3t}\right)^{\alpha'+j} - \frac{\alpha'+j}{(\beta+2t)} \left(\frac{\beta+t}{\beta+2t}\right)^{\alpha'+j} \left(\frac{\beta+t}{\beta+2t}\right)^{\alpha'+j} \right] \right\} \\
& + [Var[k|\alpha', \alpha'', \beta, t, w]] \{ [\alpha' / (\beta+2t) (\beta+t/\beta+2t)^{\alpha'}]^2 + [\alpha'' / \beta (\beta+t/\beta+2t)^{\alpha'}]^2 \}.
\end{aligned}$$

Prova. Seja $v > 0$. Analogamente a prova da Proposição 3.5.2 (i),
 $P[V(t) \leq v | \alpha', \alpha'', \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots]$

$$= \sum_{k \geq w} \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]}$$

$$P[V(t) \leq v | k; \alpha', \alpha'', \beta, t, ; X_1(t) = x_1, \dots, X_k(t) = x_k] \pi(k | \alpha', \beta, t, w),$$

$v > 0$. Então,

$$E[V(t) | \alpha', \alpha'', \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots]$$

$$= \sum_{k \geq w} \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]}$$

$$E[V(t) | k; \alpha', \alpha'', \beta, t, ; X_1(t) = x_1, \dots, X_k(t) = x_k] \pi(k | \alpha', \beta, t, w),$$

e

$$Var[V(t) | \alpha', \alpha'', \beta, t, A_1(t) = a_1, A_2(t) = a_2, \dots]$$

$$= \sum_{k \geq w} \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots, k]}$$

$$\text{Var}[V(t)|k; \alpha', \alpha'', \beta, t, ; X_1(t) = x_1, \dots, X_k(t) = x_k] \pi(k|\alpha', \beta, t, w).$$

O Lema 3.5.1 (iii) completa a prova de (i) e (ii) acima.

3.6 Resultados assintóticos

Para toda distribuição a priori π consideremos a função $\tau : \mathcal{N} \rightarrow \mathcal{N}$, definida por

$$\tau(x) = \begin{cases} \inf A_{\pi, x}, & \text{se } A_{\pi, x} \neq \emptyset, \\ \sup\{n \in \mathcal{N}^*; \pi(n) > 0\}, & \text{se } A_{\pi, x} = \emptyset, \end{cases}$$

para todo $x \in \mathcal{N}$.

Dados $k \in \mathcal{N}^*$, $\lambda_1, \dots, \lambda_k$, $\lambda_i > 0$, $i = 1, 2, \dots, k$ e $q_1, \dots, q_k$, $0 < q_i \leq 1$, $i = 1, 2, \dots, k$, seja P a medida de probabilidade $P[. | k, \lambda_1, \dots, \lambda_k, q_1, \dots, q_k]$ da superposição dos k processos de Poisson com taxas $q_1 \lambda_1, \dots, q_k \lambda_k$.

Pelos resultados assintóticos do capítulo 2, seção 2.6 na página 36, vimos que, quando $t \rightarrow \infty$,

$$E[k^r | \alpha', \beta, t, W_t] \xrightarrow{q.c.[P]} (\tau(k))^r,$$

$r \in \mathcal{N}^*$, e que

$$\text{Var}[k | \alpha', \beta, t, W_t] \xrightarrow{q.c.[P]} 0.$$

Vimos também no capítulo 2 que os estimadores dados pelas esperanças a posteriori de $U(t)$ e $V(t)$, e pelas variâncias a posteriori de $U(t)$ e $V(t)$ convergem quase certamente para zero, quando $t \rightarrow \infty$. Nesta seção mostramos que nesta versão do problema, no qual há erros gerados que não serão detectados, estes estimadores podem não convergir para zero.

Teorema 3.6.1 *Seja π uma distribuição a priori para k tal que $A_{\pi,k} \neq \emptyset$. $\alpha' > 0$, $\alpha'' > 0$ e $\beta > 0$. Então, quando $t \rightarrow \infty$,*

$$(i) E[U(t)|\alpha', \alpha'', \beta, t, W_t] \xrightarrow{q.c.[P]} (\tau(k) - k) \frac{\alpha''}{\beta};$$

$$(ii) Var[U(t)|\alpha', \alpha'', \beta, t, W_t] \xrightarrow{q.c.[P]} (\tau(k) - k) \frac{\alpha''}{\beta^2};$$

$$(iii) E[V(t)|\alpha', \alpha'', \beta, t, A_1(t), A_2(t), \dots] \xrightarrow{q.c.[P]} (\tau(k) - k) \frac{\alpha''}{\beta} \left(\frac{1}{2}\right)^{\alpha'};$$

$$(iv) Var[V(t)|\alpha', \alpha'', \beta, t, A_1(t), A_2(t), \dots] \xrightarrow{q.c.[P]}$$

$$(\tau(k) - k) \left\{ \left(\frac{\alpha''}{\beta} \right)^2 \left[\left(\frac{1}{3} \right)^{\alpha'} - \left(\frac{1}{4} \right)^{\alpha'} \right] + \left(\frac{\alpha''}{\beta^2} \right) \left(\frac{1}{3} \right)^{\alpha'} \right\}.$$

Prova. Do Teorema 2.6.3 (i) temos que, quando $t \rightarrow \infty$, $E[k|\alpha', \beta, t, W_t] \xrightarrow{q.c.[P]} \tau(k)$, e do Lema 2.6.2 (i) temos que $W_t \xrightarrow{q.c.[P]} k$. Portanto, da Proposição 3.5.2 (ii) e (iii), provamos (i) e (ii). Para provarmos (iii) e (iv) observemos as expressões de $E[V(t)|\alpha', \alpha'', \beta, t, a_1, a_2, \dots]$ e $Var[V(t)|\alpha', \alpha'', \beta, t, a_1, a_2, \dots]$ na Proposição 3.5.3 (i) e (ii). Notamos que, quando $t \rightarrow \infty$, $E[k|\alpha', \beta, t, W_t] \xrightarrow{q.c.[P]} \tau(k)$, $W_t \xrightarrow{q.c.[P]} k$ e $A_j(t) \xrightarrow{q.c.[P]} 0$, desde que $A_j(t) = \sum_{i=1}^k I_{(X_i(t)=j)}$ e, para $i = 1, 2, \dots, k$, $X_i(t) \xrightarrow{q.c.[P]} \infty$, $j = 1, 2, \dots$. $\diamond$

Se $\pi(k) > 0$ e $\tau(k) = k$, todas as quantidades acima convergem para zero. Se $\pi(k) = 0$ e $\tau(k) > k$, as quantidades acima não convergem para zero, ou seja, neste caso, mesmo assintoticamente as estimativas para a taxa de falha restante do sistema revisado não atingem zero.

Capítulo 4

Estimação bayesiana do número de espécies de uma população

4.1 Introdução

Consideremos uma população particionada em k espécies $e_1, e_2, \dots, e_k$, $k \in \mathcal{N}^*$, e a seleção sequencial de elementos desta população como um processo multinomial tal que $P\{e_1\} = p_1$, $P\{e_2\} = p_2$, $\dots$, $P\{e_k\} = p_k$, $p_i \geq 0$, $i = 1, 2, \dots, k$, $\sum_{i=1}^k p_i = 1$. Para $n = 1, 2, \dots$ seja $X_i(n)$ o número de ocorrências de e_i numa amostra de tamanho n , $i = 1, 2, \dots, k$. Para todo $n \in \mathcal{N}^*$, $(X_1(n), X_2(n), \dots, X_k(n))$ tem distribuição multinomial com parâmetros $(n; p_1, p_2, \dots, p_k)$. Definamos

$$A_r(n) = \sum_{i=1}^k I_{(X_i(n)=r)}, r = 1, 2, \dots, \quad (4.1)$$

ou seja, $A_r(n)$ é o número de espécies observadas exatamente r vezes na amostra, $r = 1, 2, \dots$. A quantidade $A_r(n)$ é conhecida como frequência da frequência r e $(A_1(n), A_2(n), \dots)$ é conhecido como o vetor frequência das frequências. Observemos que

$$\sum_{r \geq 1} r A_r(n) = \sum_{r \geq 1} r \sum_{i=1}^k I_{(X_i(n)=r)} = \sum_{i=1}^k X_i(n) = n$$

é o tamanho da amostra.

$$W(n) = \sum_{r \geq 1} A_r(n) = \sum_{r \geq 1} \sum_{i=1}^k I_{(X_i(n)=r)} = \sum_{i=1}^k I_{(X_i(n)>0)}$$

é o número de espécies distintas observadas na amostra e

$$A_0(n) = \sum_{i=1}^k I_{(X_i(n)=0)} = k - W(n) \quad (4.2)$$

é o número de espécies que não foram observadas na amostra.

No problema das espécies, além de k , e $p_1, p_2, \dots, p_k$ serem desconhecidos, as espécies $e_1, e_2, \dots, e_k$ não são identificáveis no processo de observação da amostragem multinomial. Portanto, $X_1(n), X_2(n), \dots, X_k(n)$ não são observáveis; por outro lado, a partição da amostra induzida pelas diferentes espécies representadas na amostra é observável, isto é, as quantidades aleatórias $A_r(n)$, $r = 1, 2, \dots$ são observáveis. Observemos que $A_0(n) = k - W(n)$ não é observável. As quantidades de interesse sob esta formulação são:

k , número de espécies distintas na população,

$U(n) = \sum_{i=1}^k p_i I_{(X_i(n)=0)}$, probabilidade condicional de se descobrir uma nova espécie em uma observação adicional, e

$V(n) = E[U(n)] = \sum_{i=1}^k p_i (1 - p_i)^n$ a correspondente probabilidade incondicional de se descobrir uma nova espécie.

Com base em $A_1(n), A_2(n), \dots$, quer-se estimar k , $U(n)$ e $V(n)$. A literatura existente sobre a estimação destas quantidades é, em grande parte, frequentista. Good [28] propõe $\frac{A_1(n)}{n}$ como um estimador para $U(n)$. Robbins [56] propõe $\frac{A_1(n+1)}{n+1}$ como um estimador para $U(n)$ e observa que ele é "não viciado" para $U(n)$, no sentido de que $E[\frac{A_1(n+1)}{n+1}] = E[U(n)]$. Starr [59] considera o problema de estimação de $V(n)$ e

conjectura que $\frac{A_1(n+1)}{n+1}$ é um estimador não viciado de variância uniformemente mínima para $V(n)$, conjectura esta provada por Clayton e Frees [13]. Cohen e Sackrowitz [14] obtêm resultados que caracterizam, sob o ponto de vista da teoria da decisão, $\frac{A_1(n)}{n}$ e $\frac{A_1(n+1)}{n+1}$ como estimadores de $U(n)$ e $V(n)$, respectivamente. Eles mostram que se o número total de espécies for finito, $\frac{A_1(n)}{n}$ e $\frac{A_1(n+1)}{n+1}$ são estimadores inadmissíveis para $U(n)$, segundo a perda quadrática; se o número de espécies for infinito eles são admissíveis, segundo a perda quadrática. Enquanto $\frac{A_1(n+1)}{n+1}$ é um estimador não viciado de variância uniformemente mínima para $V(n)$, $\frac{A_1(n+1)}{n+1}$ não é um estimador não viciado de erro quadrático médio uniformemente mínimo para $U(n)$; tal estimador para $U(n)$ não existe.

A estimação de k é um problema de difícil solução (veja Bunge e Fitzpatrick [4]). Muitas tentativas foram feitas para solucionar o problema das espécies. As primeiras tentativas foram feitas sob a suposição de que $p_1 = p_2 = \dots = p_k = \frac{1}{k}$. O estimador de máxima verossimilhança para k é aproximadamente igual a solução da equação

$$W(n) = k(1 - e^{-\frac{n}{k}}), \quad (4.3)$$

(veja Good [27, p.73]). Outras aproximações para o estimador de máxima verossimilhança são dadas por Darroch [16], Driml e Ullrich [17], Johnson e Kotz [38], McNeil [48], Holst [36], Ivchenko e Timonina [37], Esty [23] e Arnold e Beaver [1]. Darroch [16] mostra também que se $n \geq k$, o estimador não viciado de variância uniformemente mínima para k é dado por

$$\frac{S(n+1, W(n))}{S(n, W(n))}, \quad (4.4)$$

onde $S(n, W(n))$ é o número de Stirling de segunda espécie (veja Charalambides e Singh [12]). Nayak e Christman [53] mostram que este

estimador de k decresce estocasticamente quando as probabilidades $p_1, p_2, \dots, p_k$ vão se tornando mais heterogêneas (no sentido de majoração), o que implica que este estimador é não viciado somente quando $p_1 = p_2 = \dots = p_k = \frac{1}{k}$.

Se nenhuma suposição for feita sobre $p_1, p_2, \dots, p_k$, os métodos padrões de estimação não funcionam. Neste caso um estimador não viciado para k não existe (veja Engen [21, p.28]) e o estimador de máxima verossimilhança de k é infinito (veja Boender e Rinnooy Kan [2]). Na busca de um estimador para k que considere a heterogeneidade das proporções populacionais, Chao [7] usa a relação em (4.2), $k = W(n) + A_0(n)$, obtém uma relação entre $E[A_0(n)]$ e $E[A_r(n)]$, $r = 1, 2, \dots$, substitui $E[A_r(n)]$ por $A_r(n)$, para $r = 1, 2, \dots$ e obtém

$$W(n) + \frac{A_1(n)^2}{2A_2(n)} \quad (4.5)$$

como estimador de k . Posteriormente, Chao e Lee [9] mostram que $\frac{E[W(n)]}{E[1-U(n)]} = k - \frac{A_1(n)}{E[1-U(n)]}\gamma^2$, onde γ é o coeficiente de variação das proporções populacionais, substituem $E[U(n)]$ pela estimativa de Good $\frac{A_1(n)}{n}$ e γ por um estimador $\hat{\gamma}$, que depende de $A_r(n)$, $r = 1, 2, \dots$ e obtém

$$\frac{W(n)}{1 - \frac{A_1(n)}{n}} + \frac{\frac{nA_1(n)}{n}}{1 - \frac{A_1(n)}{n}}\hat{\gamma}^2, \quad (4.6)$$

como estimador de k . Observemos que, enquanto (4.3) e (4.4) dependem de $A_1(n), A_2(n), \dots$ somente através de $W(n)$, número de espécies distintas observadas na amostra, (4.5) e (4.6) dependem de $A_1(n), A_2(n), \dots$ e consideram a heterogeneidade das frequências das espécies na amostra. Chao, Yip e Huey-Shyan [10, 1996] desenvolvem uma classe de estimadores via teoria de Martingais que inclui o estimador de máxima verossimilhança para o caso equiprovável e o estimador (4.6). Até o momento, não encontramos na literatura fortes recomendações para

o uso de qualquer um destes estimadores para k . O problema das espécies é ainda considerado sem solução satisfatória (veja Bunge e Fitzpatrick [4]).

Abordagens bayesianas deste problema não são tão frequentes na literatura. A distribuição a posteriori de k , com exceção de casos muito particulares, não tem uma forma explícita e boa parte da literatura bayesiana tem se concentrado na obtenção de suas estatísticas descritivas como média, variância e moda. Hill [35] considera a amostragem hipergeométrica de uma população finita de tamanho N , supõe que, dado k , as k frequências absolutas de cada classe da população tem distribuição uniforme no conjunto $\{(f_1, f_2, \dots, f_k) : \sum_{i=1}^k f_i = N, f_i \geq 1, i = 1, 2, \dots, k\}$ e que k tem distribuição binomial negativa truncada, obtendo a média e a variância a posteriori de k . Lewins e Joanes [44] supõem que, dado k , $(p_1, p_2, \dots, p_k)$ tem distribuição Dirichlet simétrica com parâmetro α e que k tem distribuição binomial negativa truncada em zero. Eles verificam que se $\alpha = 1$, a distribuição a posteriori para k é o limite da distribuição a posteriori obtida por Hill [35], quando o tamanho da população tende a infinito. Boender e Rinnooy Kan [3] consideram a amostragem multinomial, supõem que, dado k , as proporções populacionais têm distribuição Dirichlet simétrica com parâmetro α e consideram uma priori imprópria uniforme em $\mathcal{N}^*$ para k . Eles verificam numericamente que a média e a moda da distribuição a posteriori são muito sensíveis ao valor de α . Leite e Bolfarine [42] consideram a amostragem multinomial no caso equiprovável $p_1 = p_2 = \dots = p_k = \frac{1}{k}$, supõem uma priori qualquer para k e estudam analiticamente propriedades monotônicas da média a posteriori de k .

Neste capítulo, estudamos propriedades relativas à análise bayesiana deste modelo. Supomos que, dado k , as proporções populacionais têm distribuição Dirichlet simétrica com parâmetro α , e que k tem uma distribuição de probabilidade qualquer ou imprópria uniforme

nos inteiros positivos. Esta classe de prioris contém as prioris referidas acima e que vêm sendo adotadas na literatura. Na seção 4.2 apresentamos a função de verossimilhança de $(k, p_1, p_2, \dots, p_k)$ dados $A_1(n) = a_1, A_2(n) = a_2, \dots$, como uma soma de termos multinomiais. Na seção 4.3 descrevemos a estrutura da distribuição a priori. Na seção 4.4 apresentamos várias características deste modelo bayesiano. Em (4.12) observamos que a função de verossimilhança de k depende dos dados $a_1, a_2, \dots$ somente através de $w = \sum_{r \geq 1} a_r$, número de espécies distintas observadas. Portanto, como em (4.3) e (4.4), as estimativas de Bayes dependem da amostra de espécies somente através de w , não levando em conta a heterogeneidade das frequências das frequências das espécies a_r , $r = 1, 2, \dots$. Mostramos também que, se o tamanho da amostra for maior que o número de espécies observadas na amostra em pelo menos $r + 2$, a média a posteriori de k^r está bem definida, mesmo que a priori para k seja imprópria uniforme em $\mathcal{N}^*$. Na seção 4.5 apresentamos os resultados numéricos obtidos por Boender e Rinnooy Kan [3] para ilustrar o fato de que, neste modelo bayesiano o valor de α reflete a abundância relativa de espécies na população, com valores grandes de α produzindo pequenas estimativas para k , e valores pequenos de α produzindo grandes estimativas para k . A distribuição a posteriori de k depende de três parâmetros: α , tamanho da amostra, n , e número de espécies distintas observadas na amostra, w . Na seção 4.6 provamos que a distribuição a posteriori de k é estocasticamente decrescente em α , para n e w fixados. Assim obtemos uma justificativa analítica para a papel de α neste modelo como a abundância relativa de espécies na população. Mostramos também que a distribuição a posteriori de k é estocasticamente crescente em w , para α e n fixados, e estocasticamente decrescente em n , para α e w fixados. A média a posteriori de k também depende de α , n e w ; estas ordenações estocásticas implicam que a média a posteriori de k constitui-se em uma sequência

não crescente em α , não decrescente em w e não crescente em n . As distribuições a posteriori de $U(n)$ e $V(n)$ não têm formas explícitas. Na seção 4.7 mostramos que a distribuição de $U(n)$ é uma mistura de distribuições Beta, onde a função misturadora é a distribuição a posteriori de k . Obtemos expressões para a média e a variância a posteriori de $U(n)$ e $V(n)$ e veremos que a distribuição a posteriori de $U(n)$ também depende dos dados somente através de w , enquanto que a distribuição a posteriori de $V(n)$ depende das frequências das frequências $A_r(n) = a_r$, $r = 1, 2, \dots$. Na seção 4.8 provamos que a média a posteriori de k converge quase certamente para k e que a variância a posteriori de k , as médias e variâncias a posteriori de $U(n)$ e $V(n)$ convergem quase certamente para zero.

4.2 O modelo estatístico e a função de verossimilhança

Dados $k, p_1, p_2, \dots, p_k$, a distribuição conjunta de $X_1(n), X_2(n), \dots, X_k(n)$ é dada por

$$\begin{aligned} P[X_1(n) = x_1, X_2(n) = x_2, \dots, X_k(n) = x_k | k, p_1, p_2, \dots, p_k] \\ = \binom{n}{x_1, x_2, \dots, x_k} \prod_{i=1}^k p_i^{x_i} \end{aligned} \quad (4.7)$$

para quaisquer $x_1, x_2, \dots, x_k \in \mathcal{N}$ tais que $\sum_{i=1}^k x_i = n$. Assim, a distribuição das quantidades observáveis $(A_r(n), r = 1, 2, \dots)$ definidas em (4.1) é uma soma de termos como em (4.7) e, para quaisquer $a_1, a_2, \dots \in \mathcal{N}$ tais que $\sum_{r \geq 1} r a_r = n$ e $\sum_{r \geq 1} a_r \leq k$,

$$\begin{aligned} P[A_1(n) = a_1, A_2(n) = a_2, \dots, | k, p_1, p_2, \dots, p_k] \\ = \sum_{(x_1, x_2, \dots, x_k) \in [a_1, a_2, \dots, k]} \binom{n}{x_1, x_2, \dots, x_k} \prod_{i=1}^k p_i^{x_i}, \end{aligned} \quad (4.8)$$

onde

$$[a_1, a_2, \dots; k] = \{(y_1, y_2, \dots, y_k) \in \mathcal{N}^k : \sum_{i=1}^k I_{\{r\}}(y_i) = a_r, r = 1, 2, \dots\}.$$

A cardinalidade de $[a_1, a_2, \dots; k]$ é

$$\#[a_1, a_2, \dots; k] = \binom{k}{w} \binom{w}{a_1, a_2, \dots},$$

onde $w = \sum_{r \geq 1} a_r$ é o número de espécies distintas observadas na amostra.

Sejam $a_1, a_2, \dots \in \mathcal{N}$ tais que $\sum_{r \geq 1} r a_r = n$. A função de verossimilhança de $(k, p_1, p_2, \dots, p_k)$, dados $a_1, a_2, \dots$, é dada por

$$L_{(a_1, a_2, \dots)}(k, p_1, p_2, \dots, p_k) = \sum_{(x_1, x_2, \dots, x_k) \in [a_1, a_2, \dots; k]} \binom{n}{x_1, x_2, \dots, x_k} \prod_{i=1}^k p_i^{x_i} I_{\Delta_k}(p_1, p_2, \dots, p_k) I_{\{w, w+1, \dots\}}(k), \quad (4.9)$$

onde $\Delta_k = \{(p_1, p_2, \dots, p_k) : p_i \geq 0, i = 1, \dots, k, \sum_{i=1}^k p_i = 1\}$. Dados $a_1, a_2, \dots$ queremos estimar k , número de espécies na população, $U(n) = \sum_{i=1}^k p_i I_{(X_i(n)=0)}$, probabilidade condicional de se descobrir uma nova espécie em uma observação adicional e $V(n) = E[U(n) | k, p_1, p_2, \dots, p_k] = \sum_{i=1}^k p_i (1 - p_i)^n$, a correspondente probabilidade incondicional de se descobrir uma nova espécie.

4.3 A estrutura da priori

Seja $\{\pi(k), k \in \mathcal{N}^*\}$ uma função de probabilidade qualquer definida no conjunto $\mathcal{N}^*$ e seja $\alpha > 0$. Suponhamos que, dados $k \in \mathcal{N}^*$ e α , $(p_1, p_2, \dots, p_k)$ tem distribuição Dirichlet simétrica com parâmetro α e que, dado α , a distribuição de k não depende de α e é dada por

$\{\pi(k), k \in \mathcal{N}^*\}$. Então a distribuição conjunta de $(p_1, p_2, \dots, p_k, k)$, dado α , é

$$\begin{aligned} \pi(p_1, p_2, \dots, p_k, k | \alpha) &= \pi(p_1, p_2, \dots, p_k | k, \alpha) \pi(k) \\ &= \frac{\Gamma(k\alpha)}{\{\Gamma(\alpha)\}^k} \prod_{i=1}^k p_i^{\alpha-1} \pi(k) I_{\Delta_k}(p_1, p_2, \dots, p_k) I_{\mathcal{N}^*}(k). \end{aligned} \quad (4.10)$$

Boender e Rinnooy Kan [3] observam que se $\alpha = 1$, a distribuição de $(p_1, p_2, \dots, p_k)$, dado k , é igual a distribuição uniforme no simplex Δ_k . Para $0 < \alpha < 1$ a distribuição atinge seu mínimo no ponto $(1/k, 1/k, \dots, 1/k)$, centro do simplex Δ_k ; quando α decresce o mínimo no centro de Δ_k decresce e a distribuição de probabilidade vai se concentrando nos cantos de Δ_k . Se $\alpha \neq 1$ a distribuição é simétrica em torno do centro de Δ_k . Para $\alpha > 1$ a distribuição de probabilidade atinge seu único ponto de máximo no centro de Δ_k e este máximo vai aumentando, na medida em que α vai aumentando: se $\alpha = \infty$ as k proporções populacionais são iguais a $\frac{1}{k}$ com probabilidade 1. Observemos que o caso equiprovável é um caso particular desta priori que estamos considerando.

4.4 Características do modelo bayesiano

Proposição 4.4.1 *As suposições acima implicam as seguintes propriedades:*

- (i) dados k, α e $X_1(n) = x_1, X_2(n) = x_2, \dots, X_k(n) = x_k, (p_1, p_2, \dots, p_k)$ tem distribuição Dirichlet com parâmetros $\alpha + x_1, \alpha + x_2, \dots, \alpha + x_k$;
- (ii) dados k, α e $X_1(n) = x_1, X_2(n) = x_2, \dots, X_k(n) = x_k$, tais que $\sum_{i=1}^k x_i = n$ e $w = \sum_{i=1}^k I_{\mathcal{N}^*}(x_i) < k$, $U(n) = \sum_{i=1}^k p_i I_{(X_i(n)=0)}$

é uma variável aleatória com distribuição Beta com parâmetros $(k - w)\alpha$ e $w\alpha + n$:

(iii) dados k e α , $(X_1(n), X_2(n), \dots, X_k(n))$ tem distribuição Dirichlet-Multinomial simétrica com parâmetros n e α , ou seja, para quaisquer $x_1, x_2, \dots, x_k \in \mathcal{N}$ tais que $\sum_{i=1}^k x_i = n$,

$$\begin{aligned} P[X_1(n) = x_1, X_2(n) = x_2, \dots, X_k(n) = x_k | k, \alpha] \\ = n! \frac{\Gamma(k\alpha)}{\Gamma(k\alpha + n)} \prod_{i=1}^k \frac{\Gamma(\alpha + x_i)}{x_i! \Gamma(\alpha)}; \end{aligned}$$

(iv) dados k e α , $W(n)$ tem distribuição dada por

$$\begin{aligned} P[W(n) = w | k, \alpha] \\ = \binom{k}{w} \frac{n! \Gamma(k\alpha)}{\Gamma(k\alpha + n)} \sum_{m=1}^w (-1)^{w-m} \binom{w}{m} \frac{\Gamma(m\alpha + n)}{\Gamma(m\alpha) n!}, \end{aligned}$$

para $w = 1, \dots, \min\{n, k\}$;

(v) dados k e α , a distribuição de $(A_1(n), A_2(n), \dots)$ é dada por

$$\begin{aligned} P[A_1(n) = a_1, A_2(n) = a_2, \dots | k, \alpha] \\ = \binom{k}{w} \binom{w}{a_1, a_2, \dots} \frac{n! \Gamma(k\alpha)}{\Gamma(k\alpha + n)} \prod_{r \geq 1} \left(\frac{\Gamma(\alpha + r)}{r! \Gamma(\alpha)} \right)^{a_r}, \end{aligned} \quad (4.11)$$

para quaisquer $a_1, a_2, \dots$ tais que $a_r \in \mathcal{N}$, $r = 1, 2, \dots$, $\sum_{r \geq 1} r a_r = n$ e $w = \sum_{r \geq 1} a_r \leq k$.

Prova: (i) segue do fato que a distribuição Dirichlet está naturalmente conjugada com a distribuição multinomial; (ii) segue da seguinte propriedade da distribuição Dirichlet: se $\gamma_1, \gamma_2, \dots, \gamma_l$ são inteiros tais que $0 < \gamma_1 \leq \gamma_2 \leq \dots \leq \gamma_l \leq k$, dados k e $X_1(n) = x_1, X_2(n) = x_2, \dots, X_k(n) = x_k$, $\sum_{j=1}^l p_{\gamma_j}$ tem distribuição Beta com parâmetros

$\sum_{j=1}^l (\alpha + x_{\gamma_j})$ e $k\alpha + n - \sum_{j=1}^l (\alpha + x_{\gamma_j})$; (iii) segue da definição da distribuição Dirichlet-Multinomial; (iv) Emigh [19] mostra que, dados k e $(p_1, p_2, \dots, p_k)$,

$$P[W(n) = w | k, p_1, p_2, \dots, p_k] = \sum_{m=1}^w \binom{k-m}{w-m} (-1)^{w-m} \sum_{\gamma_1 > \dots > \gamma_m} \left(\sum_{j=1}^m p_{\gamma_j} \right)^n,$$

para $w = 1, \dots, \min\{n, k\}$. Usando o fato de que, para todo $\gamma_1 > \dots > \gamma_m$, $\sum_{j=1}^m p_{\gamma_j}$ tem distribuição Beta com parâmetros $m\alpha$ e $(k-m)\alpha$, temos que

$$\begin{aligned} P[W(n) = w | k, \alpha] &= \sum_{m=1}^w \binom{k-m}{w-m} (-1)^{w-m} \sum_{\gamma_1 > \dots > \gamma_m} E\left[\left(\sum_{j=1}^m p_{\gamma_j}\right)^n\right] \\ &= \sum_{m=1}^w \binom{k-m}{w-m} (-1)^{w-m} \binom{k}{m} \frac{\Gamma(m\alpha + n) \Gamma(k\alpha)}{\Gamma(k\alpha + n) \Gamma(m\alpha)} \\ &= \binom{k}{w} \frac{n! \Gamma(k\alpha)}{\Gamma(k\alpha + n)} \sum_{m=1}^w (-1)^{w-m} \binom{w}{m} \frac{\Gamma(m\alpha + n)}{\Gamma(m\alpha) n!}; \end{aligned}$$

(v) segue do fato de que $P[A_1(n) = a_1, A_2(n) = a_2, \dots | k, \alpha]$ é igual a soma de $\binom{k}{w} \binom{w}{a_1, a_2, \dots}$ termos Dirichlet-Multinomiais de (iii). $\diamond$

Sejam $a_1, a_2, \dots \in \mathcal{N}$ tais que $\sum_{r \geq 1} r a_r = n$. Do item (v) da Proposição 4.4.1 segue que o *kernel* da função de verossimilhança de k , dado $(a_1, a_2, \dots)$, depende da amostra de espécies somente através de $w = \sum_{r \geq 1} a_r$ e é dado por

$$K_{\alpha, n, w}(k) = \binom{k}{w} \frac{\Gamma(k\alpha)}{\Gamma(k\alpha + n)}, \quad (4.12)$$

$k \geq w$; portanto, a distribuição a posteriori de k é dada por

$$\pi(k | \alpha, a_1, a_2, \dots) = \pi(k | \alpha, n, w) = \frac{\binom{k}{w} \frac{\Gamma(k\alpha)}{\Gamma(k\alpha + n)} \pi(k)}{C(\alpha, n, w)},$$

$k \geq w$, onde

$$C(\alpha, n, w) = \sum_{j=w}^{\infty} K_{\alpha, n, w}(j) \pi(j) = \sum_{j=w}^{\infty} \binom{j}{w} \frac{\Gamma(j\alpha)}{\Gamma(j\alpha + n)} \pi(j).$$

Observemos que, para toda função de probabilidade a priori π para k e para quaisquer $n, w \in \mathcal{N}^*$ tais que $n \geq w$, temos $C(\alpha, n, w) < \infty$.

Com efeito,

$$\begin{aligned} C(\alpha, n, w) &= \sum_{j=w}^{\infty} \frac{\binom{j}{w} \pi(j)}{\prod_{l=0}^{n-1} (j\alpha + l)} \\ &= \frac{1}{w!} \sum_{j=w}^{\infty} \frac{\prod_{r=0}^{w-1} (j-r) \pi(j)}{\prod_{l=0}^{n-1} (j\alpha + l)} \\ &\leq \frac{1}{\alpha^n w!} \sum_{j=w}^{\infty} \frac{\pi(j)}{j^{n-w}} \leq \frac{1}{\alpha^n w!} \sum_{j=1}^{\infty} \pi(j) = \frac{1}{\alpha^n w!}. \end{aligned}$$

Analogamente, se distribuição a priori π para k for imprópria uniforme em $\mathcal{N}^*$, temos, para $n \geq w + 2$,

$$\begin{aligned} C(\alpha, n, w) &= \frac{1}{w!} \sum_{j=w}^{\infty} \frac{\prod_{r=0}^{w-1} (j-r)}{\prod_{l=0}^{n-1} (\alpha j + l)} \\ &\leq \frac{1}{\alpha^n w!} \sum_{j=w}^{\infty} \frac{1}{j^{n-w}} \leq \frac{1}{\alpha^n w!} \sum_{j=w}^{\infty} \frac{1}{j^2} < \infty. \end{aligned}$$

Para toda função de probabilidade $\{\pi(k), k \in \mathcal{N}^*\}$ e $x \in \mathcal{N}^*$, seja

$$A_{\pi, x} = \{n \in \mathcal{N}^* : n \geq x \text{ e } \pi(n) > 0\}.$$

Logo, se o número observado de espécies distintas na amostra for w , a distribuição a posteriori de k estará bem definida se a função de probabilidade a priori π para k for tal que $A_{\pi, w} \neq \emptyset$. Para todo $\alpha > 0$, para quaisquer $n, w, r \in \mathcal{N}^*$, $n \geq w$, e para toda função de probabilidade a

priori π para k tal que $A_{\pi,w} \neq \emptyset$. denotemos por $M_r(\alpha, n, w)$ o r -ésimo momento a posteriori de k , ou seja,

$$M_r(\alpha, n, w) = \frac{\sum_{k \geq w} k^r K_{\alpha,n,w}(k) \pi(k)}{C(\alpha, n, w)}.$$

Mostramos a seguir que, se o tamanho da amostra for maior que o número observado de espécies distintas em pelo menos $r + 2$ unidades, então $M_r(\alpha, n, w) < \infty$.

Proposição 4.4.2 *Para todo $\alpha > 0$, para quaisquer $n, w, r \in \mathcal{N}^*$ tais que $n \geq w + r + 2$ e para toda priori π para k tal que $A_{\pi,w} \neq \emptyset$, temos que*

$$M_r(\alpha, n, w) < \infty.$$

A condição $n \geq w + r + 2$ implica que $M_r(\alpha, n, w) < \infty$ mesmo no caso de π ser imprópria uniforme em $\mathcal{N}^$.*

Prova: Para todo $\alpha > 0$, para quaisquer $n, w, r \in \mathcal{N}^*$ tais que $n \geq w + r + 2$ e para toda priori π para k tal que $A_{\pi,w} \neq \emptyset$, temos

$$\begin{aligned} \sum_{j \geq w} j^r K_{\alpha,n,w}(j) \pi(j) &= \sum_{j \geq w} \frac{j^r \binom{j}{w} \pi(j)}{\prod_{l=0}^{n-1} (j\alpha + l)} = \frac{1}{w!} \sum_{j \geq w} \frac{j^r \prod_{s=0}^{w-1} (j-s) \pi(j)}{\prod_{l=0}^{n-1} (j\alpha + l)} \\ &\leq \frac{1}{\alpha^n w!} \sum_{j \geq w} \frac{1}{j^{n-r-w}} \leq \frac{1}{\alpha^n w!} \sum_{j \geq w} \frac{1}{j^2} < \infty, \end{aligned}$$

o que implica o resultado. ◇

4.5 Resultados numéricos

Vamos ilustrar esta seção com alguns resultados numéricos obtidos por Boender e Rinnooy Kan [3], que usaram amostras já analisadas na literatura. As amostras estão resumidas na tabela 1 a seguir. A amostra I de espécies de borboletas foi usada por Craigh [15], a amostra II de insetos foi estudada por Fisher, Corbet e Willians [25] e a amostra III consiste de nomes de Macaulay (Bacon) descrita por Yule [68, p.63].

Boender e Rinnooy Kan [3] determinam a distribuição a posteriori de k , supondo que a priori para k é imprópria uniforme em $\mathcal{N}^*$. Seus resultados estão resumidos na tabela 1, onde podemos observar que as estimativas de k são extremamente sensíveis aos valores de α . Eles conjecturam que a escolha de valores pequenos para α (a priori as proporções populacionais não são iguais) resulta em estimativas grandes para k . Provamos na próxima seção que esta conjectura é verdadeira (Teorema 4.6.2).

Lewins e Joanes [44] supõem que k tem distribuição Binomial Negativa truncada em zero e, analisando resultados numéricos, chegam a mesma conclusão que Boender e Rinnooy Kan sobre o papel de α : "o valor de α reflete a abundância relativa das espécies na população, com α grande correspondendo a um conjunto uniforme de abundâncias relativas, enquanto valores pequenos de α refletem um conjunto muito diversificado. Portanto, se α decrescer o modelo permitirá um número maior de espécies raras, que são mais improváveis de ocorrerem na amostra."

Tabela 1. *Tamanho da amostra. n , número de classes distintas observadas na amostra. w , e estimador de máxima verossimilhança, $\hat{k}$, sob a hipótese de equiprobabilidade das classes: moda de $\pi(k|\alpha, n, w)$ no topo da cela, média de $\pi(k|\alpha, n, w)$ no meio da cela e desvio padrão de $\pi(k|\alpha, n, w)$ na parte inferior da cela.*

	$\alpha = 0.01$	$\alpha = 0.1$	$\alpha = 1$	$\alpha = 10$	$\alpha = 100$
I-Borboletas	73707	8101	1574	925	861
$n = 435, w = 341, \hat{k} = 854$	75299	8304	1608	943	877
	9171	986	169	86	78
II-Amostra de insetos	4159	555	243	240	240
$n = 15609, w = 240, \hat{k} = 240$	4182	558	243.8	240.04	240.02
	286	29	2	0.07	0.02
III-Macaulay	90308	10570	2747	2135	2069
$n = 8045, w = 2048, \hat{k} = 2092$	90400	10580	2747.9	2135.4	2069.6
	2524	264	35.5	10.2	7.4

Fonte: Boender e Rinnooy Kan [3]

4.6 Propriedades de monotonicidade estocástica da distribuição a posteriori de k

A distribuição a posteriori para k é dada por

$$\pi(k|\alpha, n, w) = \frac{\binom{k}{w} \frac{\Gamma(k\alpha)}{\Gamma(k\alpha+n)} \pi(k)}{C(\alpha, n, w)} I_{A_{\pi, w}}(k),$$

onde $\alpha > 0$, n e $w \in \mathcal{N}^*$, $n \geq w$ e π é a priori para k tal que

$$A_{\pi, w} = \{n \in \mathcal{N}^* : n \geq w \text{ e } \pi(n) > 0\} \neq \emptyset$$

e

$$C(\alpha, n, w) = \sum_{j=w}^{\infty} \binom{j}{w} \frac{\Gamma(j\alpha)}{\Gamma(j\alpha + n)} \pi(j).$$

Observemos que a distribuição a posteriori para k depende dos dados através de w e n e da priori através de π e α . Dada uma priori π para k , consideremos a família de distribuições a posteriori para k indexadas por α , n e w ,

$$\{\pi(\cdot | \alpha, n, w) : \alpha > 0, (n, w) \in R(\pi)\},$$

onde o conjunto $R(\pi)$ é definido do seguinte modo:

se π for uma função de probabilidade em $\mathcal{N}^*$ com suporte A , então

$$R(\pi) = \begin{cases} \{(x, y) \in \mathcal{N}^* \times \{1, 2, \dots, k^*\}; x \geq y, k^* = \sup(A)\}, & \text{se } A \text{ for limitado,} \\ \{(x, y) \in \mathcal{N}^* \times \mathcal{N}^*; x \geq y\}, & \text{se } A \text{ não for limitado;} \end{cases}$$

se π for imprópria uniforme em $\mathcal{N}^*$, então

$$R(\pi) = \{(x, y) \in \mathcal{N}^* \times \mathcal{N}^*; x \geq 3, y = 1, \dots, x - 2\}.$$

Observemos que $R(\pi)$ é o conjunto de dados amostrais (n, w) que, após a operação Bayesiana, produzem posteriors bem definidas para k quando a priori para k for π . Então a família acima contém somente distribuições a posteriori para k bem definidas. Vamos mostrar que esta família é estocasticamente monótona em cada um dos parâmetros α , n e w . No Teorema 4.6.2 mostramos que ela é estocasticamente decrescente em α , para n e w fixados. Lembremos que estamos denotando o r -ésimo momento a posteriori de k por $M_r(\alpha, n, w)$, para $\alpha > 0$, n, w e $r \in \mathcal{N}^*$. No Corolário 4.6.3 obtemos que $\{M_r(\alpha, n, w) : \alpha > 0, (n, w) \in R(\pi)\}$ é uma sequência não crescente em α , para n e w fixados. No Teorema 4.6.4 mostramos que a família das posteriors para k é estocasticamente crescente em w , para n e α fixados e,

consequentemente, temos no Corolário 4.6.5 que a sequência dos momentos a posteriori para k é não decrescente em w . No Teorema 4.6.6 mostramos que a posteriori para k é estocasticamente decrescente em n , para w e α fixados, e no Corolário 4.6.7 temos que os momentos a posteriori para k constituem-se em uma sequência não crescente em n , para w e α fixados.

Lema 4.6.1 *Para quaisquer $\alpha > 0$ e $j \in \mathcal{N}^*$,*

$$\frac{\partial}{\partial \alpha} \left(\frac{\Gamma(j\alpha)}{\Gamma(j\alpha + n)} \right) = \sum_{l=0}^{n-1} \frac{(-j)}{(j\alpha + l)} \frac{\Gamma(j\alpha)}{\Gamma(j\alpha + n)}. \quad (4.13)$$

Prova: Sejam $\alpha > 0$ e $j \in \mathcal{N}^*$. Observemos que

$$\frac{\partial}{\partial \alpha} \left(\frac{\Gamma(j\alpha)}{\Gamma(j\alpha + n)} \right) = \frac{(-1) \frac{\partial}{\partial \alpha} (\prod_{i=0}^{n-1} (j\alpha + i))}{(\prod_{i=0}^{n-1} (j\alpha + i))^2}, \quad (4.14)$$

e que

$$\frac{\partial}{\partial \alpha} \left(\prod_{i=0}^{n-1} (j\alpha + i) \right) = j \left(\sum_{l=0}^{n-1} \prod_{i=0, i \neq l}^{n-1} (j\alpha + i) \right). \quad (4.15)$$

Portanto, de (4.14) e (4.15) segue que

$$\frac{\partial}{\partial \alpha} \left(\frac{\Gamma(j\alpha)}{\Gamma(j\alpha + n)} \right) = \frac{\sum_{l=0}^{n-1} \frac{(-j)}{(j\alpha + l)}}{\prod_{i=0}^{n-1} (j\alpha + i)} = \sum_{l=0}^{n-1} \frac{(-j)}{(j\alpha + l)} \frac{\Gamma(j\alpha)}{\Gamma(j\alpha + n)}.$$

◇

Teorema 4.6.2 *Sejam π uma priori para k e $(n, w) \in R(\pi)$. Então*

$$0 < \alpha_1 \leq \alpha_2$$

implica que

$$P[k > x | \alpha_1, n, w] \geq P[k > x | \alpha_2, n, w], \quad -\infty < x < \infty.$$

Vale a igualdade na relação acima se π for degenerada em algum ponto de $A_{\pi, w}$.

Prova: Fixado $(n, w) \in R(\pi)$, mostramos que $P[k > x|\alpha, n, w]$ é uma função não crescente de α verificando que $\frac{\partial}{\partial \alpha}(P[k > x|\alpha, n, w])$, para todo $x \in \mathfrak{R}$, é igual a covariância de duas variáveis aleatórias negativamente correlacionadas. Seja $x \in \mathfrak{R}$ e sejam F e G funções definidas por

$$F(\alpha, n, w) = \sum_{j \geq w} I_{(x, \infty)}(j) \binom{j}{w} \frac{\Gamma(j\alpha)}{\Gamma(j\alpha + n)} \pi(j) \quad (4.16)$$

e

$$G(\alpha, n, w) = \sum_{j \geq w} \binom{j}{w} \frac{\Gamma(j\alpha)}{\Gamma(j\alpha + n)} \pi(j), \quad (4.17)$$

para todo $\alpha > 0$ e para todo $(n, w) \in R(\pi)$. As funções F e G estão bem definidas porque a distribuição a posteriori de k está bem definida. Então,

$$P[k > x|\alpha, n, w] = \frac{F(\alpha, n, w)}{G(\alpha, n, w)}$$

e

$$\frac{\partial}{\partial \alpha}(P[k > x|\alpha, n, w]) = \frac{\frac{\partial}{\partial \alpha}(F(\alpha, n, w))}{G(\alpha, n, w)} - \frac{\frac{\partial}{\partial \alpha}(G(\alpha, n, w))}{(G(\alpha, n, w))^2} F(\alpha, n, w). \quad (4.18)$$

Se $A_{\pi, w}$ for finito então, as somas em (4.16) e (4.17) são somas finitas de funções (de α). Logo, pelo Lema 4.6.1,

$$\begin{aligned} \frac{\partial(F(\alpha, n, w))}{\partial \alpha} &= \sum_{j \geq w} I_{(x, \infty)}(j) \binom{j}{w} \frac{\partial}{\partial \alpha} \left(\frac{\Gamma(j\alpha)}{\Gamma(j\alpha + n)} \right) \pi(j) \\ &= \sum_{j \geq w} I_{(x, \infty)}(j) \binom{j}{w} \sum_{l=0}^{n-1} \frac{(-j)}{(j\alpha + l)} \frac{\Gamma(j\alpha)}{\Gamma(j\alpha + n)} \pi(j) \end{aligned} \quad (4.19)$$

e

$$\frac{\partial(G(\alpha, n, w))}{\partial \alpha} = \sum_{j \geq w} \binom{j}{w} \frac{\partial}{\partial \alpha} \left(\frac{\Gamma(j\alpha)}{\Gamma(j\alpha + n)} \right) \pi(j)$$

$$= \sum_{j \geq w} \binom{j}{w} \sum_{l=0}^{n-1} \frac{(-j)}{(j\alpha + l)} \frac{\Gamma(j\alpha)}{\Gamma(j\alpha + n)} \pi(j). \quad (4.20)$$

Por outro lado, se $A_{\pi, w}$ for infinito vamos mostrar que as relações (4.19) e (4.20) também são válidas. Com efeito, se $A_{\pi, w}$ for infinito então, as somas em (4.16) e (4.17) são séries de funções (de α) convergentes. Pelo Lema 4.6.1, as séries das derivadas (em relação a α) dos termos das séries (4.16) e (4.17) são as séries de funções (4.19) e (4.20), respectivamente. Desde que $\sum_{l=0}^{n-1} \frac{j}{j\alpha + l} \leq \frac{n}{\alpha}$, $j = w, w+1, \dots$, segue pelo critério de Weierstrass, que, para todo $\alpha_0 > 0$ as séries de funções (4.19) e (4.20) são uniformemente convergentes no intervalo (α_0, ∞) . Portanto, $\frac{\partial}{\partial \alpha}(F(\alpha, n, w))$ e $\frac{\partial}{\partial \alpha}(G(\alpha, n, w))$ são iguais as séries de funções (4.19) e (4.20), respectivamente, para todo $\alpha > 0$.

Logo,

$$\frac{\frac{\partial}{\partial \alpha}(F(\alpha, n, w))}{G(\alpha, n, w)} = E[I_{(k>x)} \sum_{l=0}^{n-1} \frac{(-k)}{(k\alpha + l)} | \alpha, n, w]$$

e

$$\frac{\frac{\partial}{\partial \alpha}(G(\alpha, n, w))}{G(\alpha, n, w)} = E[\sum_{l=0}^{n-1} \frac{(-k)}{(k\alpha + l)} | \alpha, n, w]$$

e, por (4.18)

$$\frac{\partial}{\partial \alpha}(P[k > x | \alpha, n, w]) = Cov[I_{(k>x)}, \sum_{l=0}^{n-1} \frac{(-k)}{(k\alpha + l)} | \alpha, n, w].$$

Consideremos as funções g e h definidas por $g(j) = I_{(x, \infty)}(j)$ e $h(j) = \sum_{l=0}^{n-1} \frac{(-j)}{(j\alpha + l)}$, para todo $j \in \mathcal{N}^*$. Desde que g é uma função não decrescente e h é uma função decrescente então,

$$Cov[I_{(k>x)}, \sum_{l=0}^{n-1} \frac{(-k)}{(\alpha + l)} | \alpha, n, w] \leq 0,$$

o que prova o resultado. $\diamond$

Corolário 4.6.3 *Sejam π uma priori para k , $(n, w) \in R(\pi)$ e $r \in \mathcal{N}^*$. Então,*

$$0 < \alpha_1 \leq \alpha_2 \text{ implica que } M_r(\alpha_1, n, w) \geq M_r(\alpha_2, n, w).$$

Vale a igualdade na relação acima se π for degenerada em algum ponto de $A_{\pi, w}$.

Prova. A prova é direta. Basta lembrar que, se $k|\alpha_1, n, w$ é estocasticamente maior do que $k|\alpha_2, n, w$ então, para toda função crescente $f(\cdot)$, $f(k)|\alpha_1, n, w$ é estocasticamente maior que $f(k)|\alpha_2, n, w$ e tomar $f(j) = j^r$, para todo $j \in \mathcal{N}^*$. $\diamond$

Teorema 4.6.4 *Sejam π uma priori para k , $\alpha > 0$ e $(n, w+1) \in R(\pi)$. Então,*

$$P[k > x|\alpha, n, w] \leq P[k > x|\alpha, n, w+1], \quad -\infty < x < \infty.$$

Vale a igualdade na relação acima se π for degenerada em algum ponto de $A_{\pi, w+1}$.

Prova: Seja $x \in \mathfrak{R}$. Observemos que

$$E[I_{(k>x)}(k-w)|\alpha, n, w] - E[I_{(k>x)}|\alpha, n, w]E[k-w|\alpha, n, w] \geq 0, \quad (4.21)$$

desde que $I_{(k>x)}$ e $(k-w)$ são positivamente correlacionadas. Logo, como

$$\begin{aligned} E[I_{(k>x)}(k-w)|\alpha, n, w] &= \frac{\sum_{j \geq w+1} I_{(x, \infty)}(j)(j-w) \binom{j}{w} \frac{\Gamma(j\alpha)}{\Gamma(j\alpha+n)} \pi(j)}{C(\alpha, n, w)} \\ &= (w+1) \frac{\sum_{j \geq w+1} I_{(x, \infty)}(j) \binom{j}{w+1} \frac{\Gamma(j\alpha)}{\Gamma(j\alpha+n)} \pi(j)}{C(\alpha, n, w)} \\ &= (w+1) \frac{C(\alpha, n, w+1)}{C(\alpha, n, w)} \frac{\sum_{j \geq w+1} I_{(x, \infty)}(j) \binom{j}{w+1} \frac{\Gamma(j\alpha)}{\Gamma(j\alpha+n)} \pi(j)}{C(\alpha, n, w+1)} \end{aligned}$$

$$= (w+1) \frac{C(\alpha, n, w+1)}{C(\alpha, n, w)} Pr\{k > x | \alpha, n, w+1\} \quad (4.22)$$

e

$$E[k - w | \alpha, n, w] = (w+1) \frac{C(\alpha, n, w+1)}{C(\alpha, n, w)}, \quad (4.23)$$

segue de (4.21), (4.22) e (4.23) que

$$(w+1) \frac{C(\alpha, n, w+1)}{C(\alpha, n, w)} \{P[k > x | \alpha, n, w+1] - P[k > x | \alpha, n, w]\} \geq 0,$$

o que implica o resultado. $\diamond$

Corolário 4.6.5 *Sejam π uma priori para k , $\alpha > 0$, $(n, w+1) \in R(\pi)$ e $r \in \mathcal{N}^*$. Então,*

$$M_r(\alpha, n, w) \leq M_r(\alpha, n, w+1).$$

Vale a igualdade na relação acima se π for degenerada em algum ponto de $A_{\pi, w+1}$.

Prova. É a mesma do Corolário 4.6.3.

Teorema 4.6.6 *Sejam π uma priori para k , $\alpha > 0$ e $(n, w) \in R(\pi)$. Então*

$$P[k > x | \alpha, n, w] \geq P[k > x | \alpha, n+1, w], \quad -\infty < x < \infty.$$

Vale a igualdade na relação acima se π for degenerada em algum ponto de $A_{\pi, w}$.

Prova: Fixados $x \in \mathfrak{R}$, $\alpha > 0$ e $(n, w) \in R(\pi)$ sejam g e h funções definidas por

$$g(j) = \frac{1}{(j\alpha + n)} \quad \text{e} \quad h(j) = I_{(x, \infty)}(j),$$

para todo $j \in \mathcal{N}^*$. Observemos que,

$$E[g(k)h(k)|\alpha, n, w] \leq E[g(k)|\alpha, n, w]E[h(k)|\alpha, n, w], \quad (4.24)$$

desde que g é uma função decrescente de j e h é uma função não decrescente de j . Logo, como

$$\begin{aligned} E[g(k)h(k)|\alpha, n, w] &= E\left[\frac{1}{(k\alpha + n)}I_{(k>x)}|\alpha, n, w\right] \\ &= \frac{\sum_{j \geq w} \frac{I_{(x, \infty)}(j)}{(j\alpha + n)} \binom{j}{w} \frac{\Gamma(j\alpha)}{\Gamma(j\alpha + n)} \pi(j)}{C(\alpha, n, w)} \\ &= \frac{\sum_{j \geq w} I_{(x, \infty)}(j) \binom{j}{w} \frac{\Gamma(j\alpha)}{\Gamma(j\alpha + n + 1)} \pi(j)}{C(\alpha, n, w)} \\ &= \frac{C(\alpha, n + 1, w)}{C(\alpha, n, w)} \frac{\sum_{j \geq w} I_{(x, \infty)}(j) \binom{j}{w} \frac{\Gamma(j\alpha)}{\Gamma(j\alpha + n + 1)} \pi(j)}{C(\alpha, n + 1, w)} \\ &= \frac{C(\alpha, n + 1, w)}{C(\alpha, n, w)} Pr\{k > x|\alpha, n + 1, w\} \end{aligned} \quad (4.25)$$

e

$$E[g(k)|\alpha, n, w] = \frac{C(\alpha, n + 1, w)}{C(\alpha, n, w)}, \quad (4.26)$$

segue de (4.24), (4.25) e (4.26) que

$$P[k > x|\alpha, n, w] \geq P[k > x|\alpha, n + 1, w].$$

◇

Corolário 4.6.7 *Sejam π uma priori para k , $\alpha > 0$, $(n, w) \in R(\pi)$ e $r \in \mathcal{N}^*$. Então,*

$$M_r(\alpha, n, w) \geq M_r(\alpha, n + 1, w).$$

Vale a igualdade na relação acima se π for degenerada em algum ponto de $A_{\pi, w}$.

Prova. É a mesma do Corolário 4.6.3.

4.7 A estimação de $U(n)$ e $V(n)$

Proposição 4.7.1 Para todo $\alpha > 0$, para quaisquer $a_1, a_2, \dots \in \mathcal{N}$ tais que $\sum_{r \geq 1} r a_r = n$ e $w = \sum_{r \geq 1} a_r$ e para toda priori π para k tal que $(n, w) \in R(\pi)$,

$$(i) P[U(n) \leq u | \alpha, A_1(n) = a_1, A_2(n) = a_2, \dots]$$

$$= P[U(n) \leq u | \alpha, W(n) = w] = \sum_{j \geq w} B_{((j-w)\alpha; w\alpha+n)}(u) \pi(j | \alpha, n, w),$$

para todo $u > 0$, onde $B_{(0; w\alpha+n)}(u) = 1$ e, para $j > w$, $B_{((j-w)\alpha; w\alpha+n)}(\cdot)$ é a função de distribuição acumulada Beta com parâmetros $(j-w)\alpha$ e $w\alpha + n$;

$$(ii) E[U(n) | \alpha, n, w] = E[U(n) | \alpha, W(n) = w] = E\left[\frac{(k-w)\alpha}{k\alpha+n} | \alpha, n, w\right]$$

$$= \alpha(w+1) \frac{C(\alpha, n+1, w+1)}{C(\alpha, n, w)};$$

(iii) se $A_{\pi, w+1} \neq \emptyset$, então

$$\begin{aligned} & Var[U(n) | \alpha, n, w] \\ &= \alpha(w+1) \frac{C(\alpha, n+2, w+1)}{C(\alpha, n, w)} \{ \alpha E[k | \alpha, n+2, w+1] - w\alpha + 1 \} \\ & \quad - \alpha^2 (w+1)^2 \left(\frac{C(\alpha, n+1, w+1)}{C(\alpha, n, w)} \right)^2. \end{aligned}$$

Prova: (i) $P[U(n) \leq u | \alpha, A_1(n) = a_1, A_2(n) = a_2, \dots]$

$$= \sum_{k \geq w} P[U(n) \leq u | k, \alpha, A_1(n) = a_1, A_2(n) = a_2, \dots] \pi(k | \alpha, n, w)$$

$$= \sum_{k \geq w} \frac{P[U(n) \leq u, A_1(n) = a_1, A_2(n) = a_2, \dots | k, \alpha]}{P[A_1(n) = a_1, A_2(n) = a_2, \dots | k, \alpha]} \pi(k | \alpha, n, w)$$

$$\begin{aligned}
&= \pi(w|\alpha, n, w) + \sum_{k \geq w+1} \frac{P[U(n) \leq u, A_1(n) = a_1, A_2(n) = a_2, \dots | k, \alpha]}{P[A_1(n) = a_1, A_2(n) = a_2, \dots | k, \alpha]} \pi(k|\alpha, n, w) \\
&= \pi(w|\alpha, n, w) + \sum_{k \geq w+1} \left\{ \frac{\sum^* P[U(n) \leq u | k, \alpha, X_1(n) = x_1, \dots, X_k(n) = x_k]}{P[A_1(n) = a_1, A_2(n) = a_2, \dots | k, \alpha]} \right. \\
&\quad \left. \cdot P[X_1(n) = x_1, \dots, X_k(n) = x_k | k, \alpha] \pi(k|\alpha, n, w) \right\},
\end{aligned}$$

onde $\sum^* = \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots, k]}$. Então, da propriedade (ii) da Proposição 4.4.1, segue que

$$\begin{aligned}
&P[U(n) \leq u | \alpha, A_1(n) = a_1, A_2(n) = a_2, \dots] = \pi(w|\alpha, n, w) \\
&+ \sum_{k \geq w+1} \left\{ B_{((k-w)\alpha; w\alpha+n)}(u) \frac{\sum^* P[X_1(n) = x_1, \dots, X_k(n) = x_k | k, \alpha] \pi(k|\alpha, n, w)}{P[A_1(n) = a_1, A_2(n) = a_2, \dots | k, \alpha]} \right\} \\
&= \pi(w|\alpha, n, w) + \sum_{k \geq w+1} B_{((k-w)\alpha; w\alpha+n)}(u) \pi(k|\alpha, n, w) \\
&= \sum_{k \geq w} B_{((k-w)\alpha; w\alpha+n)}(u) \pi(k|\alpha, n, w), \quad u > 0;
\end{aligned}$$

(ii) segue de (i) que

$$\begin{aligned}
E[U(n)|\alpha, n, w] &= \sum_{k \geq w+1} \int_0^1 u dB_{((k-w)\alpha; w\alpha+n)}(u) \pi(k|\alpha, n, w) \\
&= \sum_{k \geq w+1} \frac{(k-w)\alpha}{k\alpha + n} \pi(k|\alpha, n, w) \\
&= \frac{1}{C(\alpha, n, w)} \sum_{k \geq w+1} (w+1) \binom{k}{w+1} \frac{\Gamma(k\alpha)}{\Gamma(k\alpha + n + 1)} \pi(k)\alpha \\
&= (w+1) \frac{C(\alpha, n+1, w+1)}{C(\alpha, n, w)} \alpha;
\end{aligned}$$

(iii) segue de (i) que

$$E[(U(n))^2 | \alpha, n, w]$$

$$\begin{aligned}
&= \sum_{k \geq w+1} \int_0^1 u^2 dB_{((k-w)\alpha; w\alpha+n)}(u) \pi(k|\alpha, n, w) \\
&= \sum_{k \geq w+1} \frac{(k-w)\alpha((k-w)\alpha+1)}{(k\alpha+n)(k\alpha+n+1)} \pi(k|\alpha, n, w) \\
&= \frac{1}{C(\alpha, n, w)} \sum_{k \geq w+1} \frac{(k-w)\alpha((k-w)\alpha+1)}{(k\alpha+n)(k\alpha+n+1)} \binom{k}{w} \frac{\Gamma(k\alpha)}{\Gamma(k\alpha+n)} \pi(k) \\
&= \frac{1}{C(\alpha, n, w)} \sum_{k \geq w+1} \alpha((k-w)\alpha+1)(w+1) \binom{k}{w+1} \frac{\Gamma(k\alpha)}{\Gamma(k\alpha+n+2)} \pi(k) \\
&= \alpha(w+1) \frac{C(\alpha, n+2, w+1)}{C(\alpha, n, w)} \{\alpha E[k|\alpha, n+2, w+1] - w\alpha + 1\}.
\end{aligned}$$

Logo, de (ii) segue o resultado. $\diamond$

Proposição 4.7.2 *Para todo $\alpha > 0$, para quaisquer $a_1, a_2, \dots \in \mathcal{N}$ tais que $\sum_{r \geq 1} r a_r = n$ e $w = \sum_{r \geq 1} a_r$ e para toda priori π para k tal que $(n, w) \in R(\pi)$,*

$$\begin{aligned}
(i) \quad &P[V(n) \leq v | \alpha, A_1(n) = a_1, A_2(n) = a_2, \dots] \\
&= \sum_{k \geq w} \frac{1}{\binom{k}{k-w, a_1, a_2, \dots}} \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]}
\end{aligned}$$

$$P[V(n) \leq v | k, \alpha, X_1(n) = x_1, X_2(n) = x_2, \dots, X_k(n) = x_k] \pi(k|\alpha, n, w),$$

para todo $v > 0$;

$$(ii) \quad E[V(n) | \alpha, A_1(n) = a_1, A_2(n) = a_2, \dots]$$

$$\begin{aligned}
&= E[(k-w) \frac{\alpha}{(n+k\alpha)} \prod_{s=0}^{n-1} (1 - \frac{\alpha+1}{n+k\alpha+s+1}) | \alpha, W(n) = w] \\
&\quad + \sum_{r \geq 1} a_r E[\frac{(r+\alpha)}{(n+k\alpha)} \prod_{s=0}^{n-1} (1 - \frac{r+\alpha+1}{n+k\alpha+s+1}) | \alpha, W(n) = w];
\end{aligned}$$

$$(iii) \quad Var[V(n) | \alpha, A_1(n) = a_1, A_2(n) = a_2, \dots]$$

$$= E[(k-w) \frac{(\alpha+1)\alpha \prod_{s=0}^{2n-1} ((k-1)\alpha + n + s)}{\prod_{s=0}^{2n+1} (k\alpha + n + s)}]$$

$$\begin{aligned}
& + \sum_{r \geq 1} a_r \frac{(\alpha + r + 1)(\alpha + r) \prod_{s=0}^{2n-1} ((k-1)\alpha + n - r + s)}{\prod_{s=0}^{2n+1} (k\alpha + n + s)} |_{\alpha, n, w} \\
& + 2 \sum_{k \geq w} \frac{1}{\binom{k}{k-w, a_1, \dots}} \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots, k]} \sum_{l=0}^n \sum_{m=0}^n \binom{n}{l} \binom{n}{m} \sum_{i=1}^k \sum_{j>i} \\
& \frac{\prod_{s=0}^{n-m} (\alpha + x_i + s) \prod_{s=0}^{n-l} (\alpha + x_j + s) \prod_{s=0}^{l+m-1} ((k-2)\alpha + n - x_i - x_j + s)}{\prod_{s=0}^{2n+1} (k\alpha + n + s)} \\
& \cdot \pi(k | \alpha, n, w) \\
& - \{ E[(k-w) \frac{\alpha}{(n+k\alpha)} \prod_{s=0}^{n-1} (1 - \frac{\alpha+1}{n+k\alpha+s+1})] |_{\alpha, n, w} \\
& + \sum_{r \geq 1} a_r E[\frac{(r+\alpha)}{(n+k\alpha)} \prod_{s=0}^{n-1} (1 - \frac{r+\alpha+1}{n+k\alpha+s+1})] |_{\alpha, n, w} \}^2.
\end{aligned}$$

Prova.

$$(i) P[V(n) \leq v | \alpha, A_1(n) = a_1, A_2(n) = a_2, \dots]$$

$$\begin{aligned}
& = \sum_{k \geq w} \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots, k]} P[V(n) \leq v | k, \alpha, X_1(n) = x_1, \dots, X_k(n) = x_k] \\
& \cdot \frac{P[X_1(n) = x_1, \dots, X_k(n) = x_k | k, \alpha]}{P[A_1(n) = a_1, A_2(n) = a_2, \dots | k, \alpha]} \pi(k | \alpha, n, w).
\end{aligned}$$

Logo, o resultado segue de (iii) e (v) da Proposição 4.4.1;

(ii) notemos, por (i), que

$$\begin{aligned}
& E[V(n) | \alpha, A_1(n) = a_1, A_2(n) = a_2, \dots] \tag{4.27} \\
& = \sum_{k \geq w} \frac{1}{\binom{k}{k-w, a_1, \dots}} \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots, k]} \\
& E[V(n) | k, \alpha, X_1(n) = x_1, \dots, X_k(n) = x_k] \pi(k | \alpha, n, w) \\
& = \sum_{k \geq w} \frac{1}{\binom{k}{k-w, a_1, \dots}} \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots, k]}
\end{aligned}$$

$$E\left[\sum_{i=1}^k p_i(1-p_i)^n | k, \alpha, X_1(n) = x_1, \dots, X_k(n) = x_k\right] \pi(k | \alpha, n, w).$$

Do item (i) da Proposição 4.4.1 segue que, para todo $(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]$,

$$\begin{aligned} & E\left[\sum_{i=1}^k p_i(1-p_i)^n | k, \alpha, X_1(n) = x_1, \dots, X_k(n) = x_k\right] \\ &= \sum_{i=1}^k \frac{\Gamma(x_i + \alpha + 1) \Gamma(n - x_i + (k-1)\alpha + n)}{\Gamma(x_i + \alpha + (n - x_i) + (k-1)\alpha + 1 + n)} \frac{\Gamma(x_i + \alpha + n - x_i + (k-1)\alpha)}{\Gamma(x_i + \alpha) \Gamma(n - x_i + (k-1)\alpha)} \\ &= \sum_{i=1}^k \frac{(x_i + \alpha)}{(n + k\alpha)} \prod_{s=0}^{n-1} \left(1 - \frac{x_i + \alpha + 1}{n + k\alpha + s + 1}\right) \\ &= (k - w) \frac{\alpha}{(n + k\alpha)} \prod_{s=0}^{n-1} \left(1 - \frac{\alpha + 1}{n + k\alpha + s + 1}\right) \\ &\quad + \sum_{r \geq 1} a_r \frac{(r + \alpha)}{(n + k\alpha)} \prod_{s=0}^{n-1} \left(1 - \frac{r + \alpha + 1}{n + k\alpha + s + 1}\right), \end{aligned}$$

o que implica, por (4.27), o resultado;

$$(iii) \text{ } Var[V(n) | \alpha, A_1(n) = a_1, A_2(n) = a_2, \dots]$$

$$= E[(V(n))^2 | \alpha, A_1(n) = a_1, A_2(n) = a_2, \dots] -$$

$$E^2[V(n) | \alpha, A_1(n) = a_1, A_2(n) = a_2, \dots].$$

Por (i) e (ii) temos

$$Var[V(n) | \alpha, A_1(n) = a_1, A_2(n) = a_2, \dots] \quad (4.28)$$

$$\begin{aligned} &= \sum_{k \geq w} \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} E[(V(n))^2 | k, \alpha, X_1(n) = x_1, \dots, X_k(n) = x_k] \\ &\quad \frac{1}{\binom{k}{k-w, a_1, \dots}} \pi(k | \alpha, n, w) \end{aligned}$$

$$\begin{aligned}
& -\{E[(k-w)\frac{\alpha}{(n+k\alpha)}\prod_{s=0}^{n-1}(1-\frac{\alpha+1}{n+k\alpha+s+1})|\alpha, n, w] \\
& + \sum_{r \geq 1} a_r E[\frac{(r+\alpha)}{(n+k\alpha)}\prod_{s=0}^{n-1}(1-\frac{r+\alpha+1}{n+k\alpha+s+1})|\alpha, n, w]\}^2.
\end{aligned}$$

Para todo k e $(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]$, temos

$$\begin{aligned}
& E[(V(n))^2|k, \alpha, X_1(n) = x_1, \dots, X_k(n) = x_k] \quad (4.29) \\
& = E[(\sum_{i=1}^k p_i(1-p_i)^n)^2|k, \alpha, x_1, \dots, x_k] \\
& = E[\sum_{i=1}^k p_i^2(1-p_i)^{2n}|k, \alpha, x_1, \dots, x_k] \\
& + 2E[\sum_{i=1}^k \sum_{j>i}^k p_i p_j (1-p_i)^n (1-p_j)^n |k, \alpha, x_1, \dots, x_k].
\end{aligned}$$

Pela Proposição 4.4.1(i), $p_i|k, \alpha, x_1, \dots, x_k$ tem distribuição Beta com parâmetros $\alpha + x_i$ e $(k-1)\alpha + n - x_i$, $i = 1, 2, \dots, k$. Então, para todo $(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]$, temos

$$\begin{aligned}
& E[\sum_{i=1}^k p_i^2(1-p_i)^{2n}|k, \alpha, x_1, \dots, x_k] \\
& = \sum_{i=1}^k \frac{\Gamma(\alpha + x_i + 2)\Gamma((k-1)\alpha + n - x_i + 2n)}{\Gamma(k\alpha + n + 2n + 2)} \frac{\Gamma(k\alpha + n)}{\Gamma(\alpha + x_i)\Gamma((k-1)\alpha + n - x_i)} \\
& = \sum_{i=1}^k \frac{(\alpha + x_i + 1)(\alpha + x_i) \prod_{s=0}^{2n-1} ((k-1)\alpha + n - x_i + s)}{\prod_{s=0}^{2n+1} (k\alpha + n + s)} \\
& = \sum_{i=1}^k \frac{(\alpha + x_i + 1)}{(k\alpha + n + 1)} \frac{(\alpha + x_i)}{(k\alpha + n)} \prod_{s=0}^{2n-1} (1 - \frac{x_i + \alpha + 2}{k\alpha + n + s + 2}) \\
& = (k-w) \frac{(\alpha + 1)}{(k\alpha + n + 1)} \frac{\alpha}{(k\alpha + n)} \prod_{s=0}^{2n-1} (1 - \frac{\alpha + 2}{k\alpha + n + 2 + s})
\end{aligned}$$

$$+ \sum_{r \geq 1} a_r \frac{(\alpha + r + 1)}{(k\alpha + n + 1)} \frac{(\alpha + r)}{(k\alpha + n)} \prod_{s=0}^{2n-1} \left(1 - \frac{r + \alpha + 2}{k\alpha + n + s + 2}\right) \quad (4.30)$$

e

$$\begin{aligned} & E\left[\sum_{i=1}^k \sum_{j>i} p_i p_j (1 - p_i)^n (1 - p_j)^n \mid k, \alpha, x_1, \dots, x_k\right] \\ &= E\left[\sum_{i=1}^k \sum_{j>i} p_i p_j (1 - p_i - p_j + p_j)^n (1 - p_j - p_i + p_i)^n \mid k, \alpha, x_1, \dots, x_k\right] \\ &= E\left[\sum_{i=1}^k \sum_{j>i} p_i p_j \sum_{l=0}^n \binom{n}{l} (1 - p_i - p_j)^l p_j^{n-l} \sum_{m=0}^n \binom{n}{m} (1 - p_j - p_i)^m p_i^{n-m} \mid k, \alpha, x_1, \dots, x_k\right] \\ &= \sum_{l=0}^n \sum_{m=0}^n \binom{n}{l} \binom{n}{m} \sum_{i=1}^k \sum_{j>i} E[p_i^{n-m+1} p_j^{n-l+1} (1 - p_i - p_j)^{l+m} \mid k, \alpha, x_1, \dots, x_k] \\ &= \sum_{l=0}^n \sum_{m=0}^n \binom{n}{l} \binom{n}{m} \sum_{i=1}^k \sum_{j>i} E\left[\left(\frac{p_i}{p_i + p_j}\right)^{n-m+1} \left(\frac{p_j}{p_i + p_j}\right)^{n-l+1} (p_i + p_j)^{2n-m-l+2} (1 - p_i - p_j)^{l+m} \mid k, \alpha, x_1, \dots, x_k\right]. \end{aligned}$$

A Proposição 4.4.1 (i) nos diz que, dados $k, \alpha, x_1, \dots, x_k, (p_1, \dots, p_k)$ tem distribuição Dirichlet com parâmetros $\alpha + x_1, \dots, \alpha + x_k$. A caracterização da distribuição Dirichlet implica, para todo $i \neq j, i, j = 1, 2, \dots, k$, que (i) $(\frac{p_i}{p_i + p_j}, \frac{p_j}{p_i + p_j})$ é independente de $(p_i + p_j, 1 - p_i - p_j)$; (ii) $(\frac{p_i}{p_i + p_j}, \frac{p_j}{p_i + p_j})$ tem distribuição Beta com parâmetros $(\alpha + x_i, \alpha + x_j)$ e (iii) $(p_i + p_j, 1 - p_i - p_j)$ tem distribuição Beta com parâmetros $(2\alpha + x_i + x_j, (k-2)\alpha + n - x_i - x_j)$. Então, para todo $(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]$, temos

$$\begin{aligned} & E\left[\sum_{i=1}^k \sum_{j>i} p_i p_j (1 - p_i)^n (1 - p_j)^n \mid k, \alpha, x_1, \dots, x_k\right] \\ &= \sum_{l=0}^n \sum_{m=0}^n \binom{n}{l} \binom{n}{m} \sum_{i=1}^k \sum_{j>i} E\left[\left(\frac{p_i}{p_i + p_j}\right)^{n-m+1} \left(\frac{p_j}{p_i + p_j}\right)^{n-l+1} \mid k, \alpha, x_1, \dots, x_k\right] \end{aligned}$$

$$\begin{aligned}
& E[(p_i + p_j)^{2n-m-l+2}(1 - p_i - p_j)^{l+m} | k, \alpha, x_1, \dots, x_k] \\
&= \sum_{l=0}^n \sum_{m=0}^n \binom{n}{l} \binom{n}{m} \sum_{i=1}^k \sum_{j>i} \frac{\Gamma(\alpha + x_i + n - m + 1) \Gamma(\alpha + x_j + n - l + 1)}{\Gamma(\alpha + x_i + \alpha + x_j + 2n + 2 - m - l)} \\
&\quad \frac{\Gamma(2\alpha + x_i + x_j)}{\Gamma(\alpha + x_i) \Gamma(\alpha + x_j)} \\
&\quad \frac{\Gamma(2\alpha + x_i + x_j + 2n + 2 - m - l) \Gamma((k-2)\alpha + n - x_i - x_j + l + m)}{\Gamma(k\alpha + n + 2n + 2)} \\
&\quad \frac{\Gamma(k\alpha + n)}{\Gamma(2\alpha + x_i + x_j) \Gamma((k-2)\alpha + n - x_i - x_j)} \\
&= \sum_{l=0}^n \sum_{m=0}^n \binom{n}{l} \binom{n}{m} \sum_{i=1}^k \sum_{j>i} \frac{\Gamma(\alpha + x_i + n - m + 1) \Gamma(\alpha + x_j + n - l + 1)}{\Gamma(\alpha + x_i) \Gamma(\alpha + x_j)} \\
&\quad \frac{\Gamma((k-2)\alpha + n - x_i - x_j + l + m)}{\Gamma(k\alpha + n + 2n + 2)} \frac{\Gamma(k\alpha + n)}{\Gamma((k-2)\alpha + n - x_i - x_j)}.
\end{aligned}$$

Então,

$$E\left[\sum_{i=1}^k \sum_{j>i} p_i p_j (1 - p_i)^n (1 - p_j)^n | k, \alpha, x_1, \dots, x_k\right] \quad (4.31)$$

$$\begin{aligned}
&= \sum_{l=0}^n \sum_{m=0}^n \binom{n}{l} \binom{n}{m} \sum_{i=1}^k \sum_{j>i} \\
&\frac{\prod_{s=0}^{n-m} (\alpha + x_i + s) \prod_{s=0}^{n-l} (\alpha + x_j + s) \prod_{s=0}^{l+m-1} ((k-2)\alpha + n - x_i - x_j + s)}{\prod_{s=0}^{2n+1} (k\alpha + n + s)}.
\end{aligned}$$

Substituindo (4.30), (4.31) em (4.29), por (4.28), segue o resultado.

◇

4.8 Resultados assintóticos

Para toda a priori π para k consideremos a função $\tau : \mathcal{N} \rightarrow \mathcal{N}$ definida por

$$\tau(x) = \begin{cases} \inf A_{\pi,x}, & \text{se } A_{\pi,x} \neq \emptyset \\ \sup\{n \in \mathcal{N}^*; \pi(n) > 0\}, & \text{se } A_{\pi,x} = \emptyset, \end{cases}$$

para todo $x \in \mathcal{N}$.

Teorema 4.8.1 *Sejam $n \in \mathcal{N}^*$, $a_1, a_2, \dots \in \mathcal{N}$ tais que $\sum_{j \geq 1} j a_j = n$ e $\sum_{j \geq 1} a_j = w$. Sejam $r \in \mathcal{N}^*$ e π uma priori para k com $(r+w)$ -ésimo momento finito tal que $(n, w) \in R(\pi)$. Então, quando $n \rightarrow \infty$,*

- (i) $M_r(\alpha, n, w) = E[k^r | \alpha, n, w] \rightarrow [\tau(w)]^r$;
- (ii) $\text{Var}[k | \alpha, n, w] \rightarrow 0$;
- (iii) $E[U(n) | \alpha, n, w] \rightarrow 0$;
- (iv) se $A_{\pi, w+1} \neq \emptyset$, então $\text{Var}[U(n) | \alpha, n, w] \rightarrow 0$;
- (v) $E[V(n) | \alpha, n, a_1, a_2, \dots] \rightarrow 0$;
- (vi) $\text{Var}[V(n) | \alpha, n, a_1, a_2, \dots] \rightarrow 0$.

Prova: Para simplificar a notação façamos $\tau(w) = \tau$. Podemos escrever

$$M_r(\alpha, n, w) = \frac{\tau^r + A_n}{1 + B_n},$$

onde

$$A_n = \sum_{j \geq \tau+1} j^r \frac{\binom{j}{w} \frac{\Gamma(j\alpha)\pi(j)}{\Gamma(j\alpha+n)}}{\binom{\tau}{w} \frac{\Gamma(\tau\alpha)\pi(\tau)}{\Gamma(\tau\alpha+n)}}$$

e

$$B_n = \sum_{j \geq \tau+1} \frac{\binom{j}{w} \frac{\Gamma(j\alpha)\pi(j)}{\Gamma(j\alpha+n)}}{\binom{\tau}{w} \frac{\Gamma(\tau\alpha)\pi(\tau)}{\Gamma(\tau\alpha+n)}}.$$

Mas,

$$A_n = \sum_{j \geq \tau+1} j^r \frac{\binom{j}{w} \pi(j)}{\binom{\tau}{w} \pi(\tau)} \prod_{i=0}^{n-1} \left(\frac{\tau\alpha + i}{j\alpha + i} \right)$$

$$\begin{aligned}
&\leq \left(\sum_{j \geq \tau+1} j^r \frac{\binom{j}{w}}{\binom{\tau}{w}} \frac{\pi(j)}{\pi(\tau)} \right) \prod_{i=0}^{n-1} \left(\frac{\tau\alpha + i}{(\tau+1)\alpha + i} \right) \\
&= \left(\sum_{j \geq \tau+1} j^r \frac{\binom{j}{w}}{\binom{\tau}{w}} \frac{\pi(j)}{\pi(\tau)} \right) \prod_{i=0}^{n-1} \left(1 - \frac{\alpha}{(\tau+1)\alpha + i} \right) \\
&\leq \left(\sum_{j \geq \tau+1} j^r \frac{\binom{j}{w}}{\binom{\tau}{w}} \frac{\pi(j)}{\pi(\tau)} \right) e^{-\alpha \sum_{i=0}^{n-1} \frac{1}{(\tau+1)\alpha + i}} \rightarrow 0,
\end{aligned}$$

quando $n \rightarrow \infty$, pois

$$\sum_{j \geq \tau+1} j^r \frac{\binom{j}{w}}{\binom{\tau}{w}} \frac{\pi(j)}{\pi(\tau)} \leq \frac{\sum_{j \geq \tau+1} j^{r+w} \pi(j)}{\pi(\tau) \binom{\tau}{w} w!} < \infty,$$

uma vez que, por hipótese, π tem $(r+w)$ -ésimo momento finito e

$$e^{-\alpha \sum_{i=0}^{n-1} \frac{1}{(\tau+1)\alpha + i}} \rightarrow 0,$$

quando $n \rightarrow \infty$. Logo, $0 \leq A_n \rightarrow 0$, quando $n \rightarrow \infty$ e, como $0 \leq B_n \leq A_n$ segue que $B_n \rightarrow 0$, quando $n \rightarrow \infty$, o que prova o item (i);

(ii) segue de (i) que

$$Var[k|\alpha, n, w] = E[k^2|\alpha, n, w] - E^2[k|\alpha, n, w] \rightarrow \tau^2 - \tau^2 = 0,$$

quando $n \rightarrow \infty$;

(iii) segue de (i) que

$$\begin{aligned}
\frac{C(\alpha, n+1, w+1)}{C(\alpha, n, w)} &= \frac{\sum_{j \geq w+1} \frac{(j-w)}{w+1} \binom{j}{w} \frac{\Gamma(j\alpha)}{\Gamma(j\alpha+n+1)} \pi(j)}{C(\alpha, n, w)} \\
&= \frac{\sum_{j \geq w+1} \frac{(j-w)}{(w+1)} \binom{j}{w} \frac{\Gamma(j\alpha)}{\Gamma(j\alpha+n)(j\alpha+n)} \pi(j)}{C(\alpha, n, w)} \leq \frac{1}{n} E[k-w|\alpha, n, w] \rightarrow 0,
\end{aligned}$$

quando $n \rightarrow \infty$. Logo, pela Proposição 4.7.1 (ii) segue o resultado.
(iv) Como

$$\frac{C(\alpha, n+2, w+1)}{C(\alpha, n, w)} \leq \frac{C(\alpha, n+1, w+1)}{C(\alpha, n, w)} \frac{1}{[(w+1)\alpha + n+1]},$$

segue, pela Proposição 4.7.1 (iii), que

$$\text{Var}[U(n)|\alpha, n, w] \leq$$

$$\frac{C(\alpha, n+1, w+1)}{C(\alpha, n, w)} \{ \alpha(E[k|\alpha, n+2, w+1] - w) + 1 \} - E^2[U(n)|\alpha, n, w] \rightarrow 0,$$

quando $n \rightarrow \infty$, desde que, por (i),

$$E[k|\alpha, n+2, w+1] \rightarrow \tau(w+1),$$

quando $n \rightarrow \infty$, por (iii), $E^2[U(n)|\alpha, n, w] \rightarrow 0$, quando $n \rightarrow \infty$ e,

$$\frac{C(\alpha, n+1, w+1)}{C(\alpha, n, w)} \rightarrow 0,$$

quando $n \rightarrow \infty$;

(v) notamos que, para quaisquer $k, p_1, \dots, p_k$,

$$\begin{aligned} V(n) &= \sum_{i=1}^k p_i (1 - p_i)^n = \frac{\sum_{i=1}^k (n+1) p_i (1 - p_i)^n}{n+1} \leq \\ &\leq \frac{k}{n+1}, \end{aligned}$$

e, portanto, temos

$$E[V(n)|\alpha, a_1, a_2, \dots] \leq \frac{E[k|\alpha, n, w]}{n+1} \rightarrow 0, \quad (4.32)$$

quando $n \rightarrow \infty$ pelo item (i) acima; o item (vi) segue diretamente de (v), pois para quaisquer $k, p_1, \dots, p_k$, $0 \leq (V(n))^2 \leq V(n) \leq 1$ e, portanto,

$$\text{Var}[V(n)|\alpha, a_1, a_2, \dots] \leq E[V(n)|\alpha, a_1, a_2, \dots] - E^2[V(n)|\alpha, a_1, a_2, \dots] \rightarrow 0,$$

quando $n \rightarrow \infty$.

◇

Dados $k \in \mathcal{N}^*$ e $p_1, p_2, \dots, p_k$ tais que $p_i > 0$, $i = 1, 2, \dots, k$ e $\sum_{i=1}^k p_i = 1$, consideremos uma sequência infinita de ensaios independentes, onde $e_1, e_2, \dots, e_k$ são os possíveis resultados (espécies) em cada ensaio e p_i a probabilidade de ocorrer e_i , $i = 1, 2, \dots, k$. Seja P a probabilidade $P[\cdot | k, p_1, p_2, \dots, p_k]$ do processo estocástico $\{(X_1(n), X_2(n), \dots, X_k(n)), n = 1, 2, \dots\}$, onde $X_i(n)$ é o número de vezes que a espécie e_i ocorre em n ensaios, $i = 1, 2, \dots, k$.

Lema 4.8.2 *Se π for uma priori para k tal que $A_{\pi,k} \neq \emptyset$, então*

- (i) $W(n) \rightarrow k$ q.c.[P];
 - (ii) $\tau(W(n)) \rightarrow \tau(k)$ q.c.[P],
- quando $n \rightarrow \infty$.

Prova: Para todo $n \in \mathcal{N}^*$, $W(n) \leq W(n+1) \leq k$ q.c.[P]. Portanto, existe uma variável aleatória, L , $L \leq k$, tal que $W(n) \rightarrow L$ q.c.[P], quando $n \rightarrow \infty$. Por outro lado, para todo $n \in \mathcal{N}^*$, $P[W(n) = k] = P[X_i(n) > 0, i = 1, 2, \dots, k] \geq 1 - \sum_{i=1}^k P[X_i(n) = 0] = 1 - \sum_{i=1}^k (1 - p_i)^n \rightarrow 1$, quando $n \rightarrow \infty$. Portanto, $W(n) \rightarrow k$ em probabilidade P , o que implica que $L = k$ q.c.[P]. O item (ii) segue diretamente de (i). ◇

Teorema 4.8.3 *Sejam $r \in \mathcal{N}^*$ e π uma priori para k com $(r+k)$ -ésimo momento finito e $A_{\pi,k} \neq \emptyset$. Então, quando $n \rightarrow \infty$,*

- (i) $M_r(\alpha, n, W(n)) = E[k^r | \alpha, n, W(n)] \rightarrow [\tau(k)]^r$ q.c.[P];
- (ii) $\text{Var}[k | \alpha, n, W(n)] \rightarrow 0$ q.c. [P];
- (iii) $E[U(n) | \alpha, n, W(n)] \rightarrow 0$ q.c. [P];
- (iv) se $A_{\pi,k+1} \neq \emptyset$, então $\text{Var}[U(n) | \alpha, n, W(n)] \rightarrow 0$, q.c. [P];
- (v) $E[V(n) | \alpha, n, A_1(n), A_2(n), \dots] \rightarrow 0$ q.c. [P];
- (vi) $\text{Var}[V(n) | \alpha, n, A_1(n), A_2(n), \dots] \rightarrow 0$ q.c. [P].

Prova: (i) Como no Teorema 4.8.1 temos que

$$M_r(\alpha, n, W(n)) = \frac{\tau(W(n))^r + A(n, W(n))}{1 + B(n, W(n))},$$

onde

$$A(n, W(n)) \leq \left[\sum_{j \geq \tau(W(n))+1} j^r \frac{\binom{j}{W(n)}}{\binom{\tau(W(n))}{W(n)}} \frac{\pi(j)}{\pi(\tau(W(n)))} \right] e^{-\alpha \sum_{i=0}^{n-1} \frac{1}{(\tau(W(n))+1)\alpha+i}} .q.c.[P]$$

Mas, como $\tau(W(n)) \leq \tau(k)$ $q.c.[P]$,

$$\sum_{j \geq \tau(W(n))+1} j^r \frac{\binom{j}{W(n)}}{\binom{\tau(W(n))}{W(n)}} \frac{\pi(j)}{\pi(\tau(W(n)))} \leq$$

$$\frac{\sum_{j \geq 1} j^{r+k} \pi(j)}{\min\{\pi(\tau(x)), x \in \{1, 2, \dots, k\}\}} < \infty \quad q.c.[P],$$

desde que, por hipótese, π tem $(r+k)$ -ésimo momento finito e

$$e^{-\alpha \sum_{i=0}^{n-1} \frac{1}{(\tau(W(n))+1)\alpha+i}} \leq e^{-\alpha \sum_{i=0}^{n-1} \frac{1}{(\tau(k)+1)\alpha+i}} \rightarrow 0,$$

quando $n \rightarrow \infty$. Logo $0 \leq A_n \rightarrow 0 q.c.[P]$, quando $n \rightarrow \infty$. Como $0 \leq B_n \leq A_n q.c.[P]$, segue que $B_n \rightarrow 0 q.c.[P]$, quando $n \rightarrow \infty$. Logo, pelo Lema 4.8.2 (ii) segue o resultado;

(ii) segue de (i) que

$$\begin{aligned} & Var[k|\alpha, n, W(n)] \\ &= E[k^2|\alpha, n, W(n)] - E^2[k|\alpha, n, W(n)] \rightarrow \tau(k)^2 - \tau(k)^2 = 0, q.c.[P], \end{aligned}$$

quando $n \rightarrow \infty$;

(iii) segue de (i) que

$$\frac{C(\alpha, n+1, W(n)+1)}{C(\alpha, n, W(n))}$$

$$\leq \frac{1}{n} E[k - W(n) | \alpha, n, W(n)] \rightarrow 0 (\tau(k) - k) = 0 \quad q.c.[P],$$

quando $n \rightarrow \infty$. Logo, pela Proposição 4.7.1 (ii) segue o resultado;
(iv) pela prova de (iv) do Teorema 4.8.1,

$$Var[U(n) | \alpha, n, W(n)] \leq$$

$$\frac{C(\alpha, n+1, W(n)+1)}{C(\alpha, n, W(n))} \{ \alpha(E[k | \alpha, n+2, W(n)+1] - W(n)) + 1 \} -$$

$$E^2[U(n) | \alpha, n, W(n)] \rightarrow 0 \quad q.c.[P],$$

quando $n \rightarrow \infty$, desde que, por (i) e pelo Lema 4.8.2 (i),

$$E[k | \alpha, n+2, W(n)+1] - W(n) \rightarrow \tau(k+1) - k,$$

quando $n \rightarrow \infty$, por (iii), $E^2[U(n) | \alpha, n, W(n)] \rightarrow 0$ e

$$\frac{C(\alpha, n+1, W(n)+1)}{C(\alpha, n, W(n))} \rightarrow 0,$$

quando $n \rightarrow \infty$; (v) e (vi) seguem diretamente de (4.32).

◇

Capítulo 5

Estimação Bayesiana do número de espécies : amostragem por captura

5.1 Introdução

Consideremos uma população com k espécies $e_1, \dots, e_k$ e um processo sequencial de um certo número de ocasiões em que, em cada ocasião um ou mais elementos da espécie e_i são capturados com probabilidade q_i ou nenhum elemento é capturado com probabilidade $1 - q_i$, independentemente dos elementos das outras espécies, onde $0 < q_i \leq 1$, $i = 1, 2, \dots, k$. Todo elemento capturado é devolvido à população e, consequentemente, as capturas nas diferentes ocasiões são independentes. Sejam $m \in \mathcal{N}^*$ e $X_i(m)$ o número de capturas da espécie e_i durante as m ocasiões, $i = 1, 2, \dots, k$. Notemos que uma espécie é capturada quando um ou mais elementos dessa espécie forem capturados. Então, $X_1(m), X_2(m), \dots, X_k(m)$ são variáveis aleatórias independentes com distribuição Binomial com parâmetros $(m, q_1), (m, q_2), \dots, (m, q_k)$, respectivamente. Definamos

$$A_r(m) = \sum_{i=1}^k I_{(X_i(m)=r)}, \quad r = 1, 2, \dots, m,$$

ou seja, $A_r(m)$ é o número de espécies capturadas exatamente r vezes nas m ocasiões, $r = 1, 2, \dots, m$. Logo,

$$\sum_{r=1}^m r A_r(m) = \sum_{r=1}^m r \sum_{i=1}^k I_{(X_i(m)=r)} = \sum_{i=1}^k X_i(m)$$

é o número total de capturas nas m ocasiões,

$$W_m = \sum_{r=1}^m A_r(m) = \sum_{r=1}^m \sum_{i=1}^k I_{(X_i(m)=r)} = \sum_{i=1}^k I_{(X_i(m)>0)}$$

é o número de espécies distintas capturadas nas m ocasiões, e

$$A_0(m) = \sum_{i=1}^k I_{(X_i(m)=0)} = k - W_m$$

é o número de espécies não capturadas.

Dados $k \in \mathcal{N}^*$ e $q_1, q_2, \dots, q_k$ tais que $q_i \in (0, 1]$, $i = 1, 2, \dots, k$, a distribuição de $A_r(m)$, $r = 1, 2, \dots, m$ é, para quaisquer $a_1, a_2, \dots, a_m \in \mathcal{N}$ tais que $\sum_{r=1}^m r a_r \leq km$ e $\sum_{r=1}^m a_r \leq k$,

$$P[A_1(m) = a_1, A_2(m) = a_2, \dots, A_m(m) = a_m | k, q_1, q_2, \dots, q_k]$$

$$= \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots, a_m; k]} \prod_{i=1}^k \binom{m}{x_i} q_i^{x_i} (1 - q_i)^{m-x_i},$$

onde

$$[a_1, a_2, \dots, a_m; k] = \{(y_1, \dots, y_k) \in \mathcal{N}^k : \sum_{i=1}^k I_{\{y_i\}}(r) = a_r, r = 1, 2, \dots, m\}.$$

A cardinalidade de $[a_1, a_2, \dots, a_m; k]$ é

$$\#[a_1, a_2, \dots, a_m; k] = \binom{k}{a_0, a_1, \dots, a_m},$$

onde $w = \sum_{r \geq 1} a_r$ é o número de espécies distintas capturadas nas m ocasiões e $a_0 = k - w$ é o número de espécies que não foram capturadas.

Então, supondo que, dado k , $q_1, q_2, \dots, q_k$ sejam variáveis aleatórias independentes e identicamente distribuídas com distribuição F segue que, para quaisquer $a_1, a_2, \dots, a_m \in \mathcal{N}$, tais que $\sum_{r=1}^m r a_r \leq km$ e $\sum_{r=1}^m a_r \leq k$,

$$\begin{aligned} P[A_1(m) = a_1, A_2(m) = a_2, \dots, A_m(m) = a_m | k, F] \\ = \binom{k}{a_0, a_1, \dots, a_m} \prod_{r=0}^m [q_r(F, m)]^{a_r}, \end{aligned}$$

onde

$$q_r(F, m) = \int_0^1 \binom{m}{r} q^r (1 - q)^{m-r} dF(q),$$

$r = 0, 1, 2, \dots, m$. Burnhan e Overton [5] introduzem este modelo, consideram k e F desconhecidos e obtêm um estimador para k pelo método jackknife. Tal estimador é uma combinação linear das frequências de capturas $A_1(m), A_2(m), \dots, A_m(m)$.

Suponhamos que a distribuição a priori para k seja π , onde π é uma função de probabilidade em $\mathcal{N}^*$ ou uma distribuição imprópria uniforme em $\mathcal{N}^*$. Dados $A_1(m) = a_1, A_2(m) = a_2, \dots, A_m(m) = a_m$, a função de verossimilhança de k é

$$L_{(a_1, a_2, \dots, a_m)}(k) = \binom{k}{a_0, a_1, \dots, a_m} \prod_{r=0}^m [q_r(F, m)]^{a_r} \propto \binom{k}{w} (q_0(F, m))^k, \quad (5.1)$$

para todo $k \geq w = \sum_{r=1}^m a_r$, onde $q_0(F, m) = \int_0^1 (1 - q)^m dF(q)$.

5.2 A distribuição a posteriori de k

Doravante, exceto na seção 5.4, quando nos referirmos a uma função de distribuição acumulada em $\mathfrak{R}$, F , fica subentendido que $F(x) = 0$, para todo $x < 0$ e $F(x) = 1$, para todo $x > 1$. Comparando (5.1) com

(2.6) na página 23. vemos que a atualização da priori de k para sua posteriori se dará como em (2.7) na página 24. Portanto, a distribuição a posteriori de k , dados $a_1, a_2, \dots, a_m \in \mathcal{N}$, é

$$\begin{aligned}\pi(k|F, m, a_1, a_2, \dots, a_m) &= \pi(k|F, m, w) \\ &= \frac{\binom{k}{w} (q_0(F, m))^k \pi(k)}{C(F, m, w)},\end{aligned}\quad (5.2)$$

$k \geq w$, onde

$$C(F, m, w) = \sum_{j \geq w} \binom{j}{w} (q_0(F, m))^j \pi(j) \quad \text{e} \quad (5.3)$$

$$q_0(F, m) = \int_0^1 (1 - q)^m dF(q).$$

A fim de que a posteriori para k esteja bem definida, π e w devem ser tais que $A_{\pi, w} = \{n \in \mathcal{N} : n \geq w \text{ e } \pi(n) > 0\} \neq \emptyset$. Como na superposição de processos independentes de Poisson, a distribuição a posteriori para k tem forma explícita para certas classes de priors:

- (i) se a priori para k for imprópria uniforme em $\mathcal{N}^*$, a posteriori para k está bem definida e é igual a distribuição de $w + k'$, onde k' tem distribuição Binomial Negativa com parâmetros $w + 1$ e $1 - q_0(F, m)$ (veja Proposição 2.4.1, página 24);
- (ii) se a priori π para k for da família exponencial tal que $A_{\pi, w} \neq \emptyset$, a distribuição a posteriori de k também pertence a família exponencial, onde a passagem da priori para a posteriori se dá como na Proposição 2.4.3, página 26, fazendo $b = q_0(F, m)$.

Vamos mostrar a seguir que os estimadores dados pelos momentos a posteriori de k são consistentes. Dados $k \in \mathcal{N}^*$ e $q_i \in (0, 1]$, $i =$

$1, \dots, k$, seja $P = P[\cdot | k, q_1, \dots, q_k]$ a distribuição de $\{(X_1(m), \dots, X_k(m)), m = 1, 2, \dots\}$, onde para todo $m = 1, 2, \dots$ $X_1(m), \dots, X_k(m)$ são independentes e $X_i(m)$ tem distribuição Binomial com parâmetros m e q_i , $i = 1, 2, \dots, k$. Analogamente ao raciocínio desenvolvido na seção 2.6, página 37, temos que $W_m \rightarrow k$ q.c. $[P]$, quando $m \rightarrow \infty$, o que implica que $\tau(W_m) \rightarrow \tau(k)$ q.c. $[P]$, quando $m \rightarrow \infty$, onde a função $\tau : \mathcal{N} \rightarrow \mathcal{N}$ é definida por

$$\tau(x) = \begin{cases} \inf A_{\pi, x}, & \text{se } A_{\pi, x} \neq \emptyset, \\ \sup\{n \in \mathcal{N}^*; \pi(n) > 0\}, & \text{se } A_{\pi, x} = \emptyset, \end{cases}$$

para todo $x \in \mathcal{N}$. Pela prova de (i) do Teorema 2.6.3, página 38, temos que

$$E[k^r | F, m, W_m] = \frac{[\tau(W_m)]^r + A(m, W_m)}{1 + B(m, W_m)},$$

$$0 \leq B(m, W_m) \leq A(m, W_m),$$

e

$$A(m, W_m) \leq \frac{(\tau(k) + r + 1)!}{\min\{\pi(\tau(x)) : 0 \leq x \leq k\}} \frac{q_0(F, m)}{[1 - q_0(F, m)]^{\tau(k) + r + 2}}, \quad \text{q.c. } [P].$$

Portanto $A(m, W_m)$ e $B(m, W_m)$ convergem para zero q.c. P , quando $m \rightarrow \infty$, pois $q_0(F, m) = \int_0^1 (1 - q)^m dF(q) \rightarrow 0$, quando $m \rightarrow \infty$. Como $[\tau(W_m)]^r \rightarrow [\tau(k)]^r$, q.c. $[P]$, temos que $E[k^r | F, m, W_m] \rightarrow [\tau(k)]^r$, q.c. $[P]$.

A seguir estudamos propriedades de ordenação da distribuição a posteriori de k .

5.3 Monotonicidade da distribuição a posteriori de k

Teorema 5.3.1 *Seja π uma distribuição a priori para k , própria ou imprópria uniforme em $\mathcal{N}^*$.*

(i) Para toda F e G funções de distribuição acumulada em $\mathfrak{R}$, $m \in \mathcal{N}^*$ e $w \in \mathcal{N}$, tal que $A_{\pi,w} \neq \emptyset$.

$$F \leq_{estoc} G$$

implica que

$$P[k > x|F, m, w] \geq P[k > x|G, m, w], \quad -\infty < x < \infty;$$

(ii) para toda função de distribuição acumulada em $\mathfrak{R}$, F , $m \in \mathcal{N}^*$ e $w \in \mathcal{N}$ tal que $A_{\pi,w+1} \neq \emptyset$,

$$P[k > x|F, m, w] \leq P[k > x|F, m, w+1] \quad -\infty < x < \infty;$$

(iii) para toda função de distribuição F acumulada em $\mathfrak{R}$, $m \in \mathcal{N}^*$ e $w \in \mathcal{N}$ tal que $A_{\pi,w} \neq \emptyset$,

$$P[k > x|F, m+1, w] \leq P[k > x|F, m, w] \quad -\infty < x < \infty.$$

Prova. Seja $x \in \mathfrak{R}$.

(i) Dados F e G tais que $F \leq_{estoc} G$, m e w , sejam g e h funções reais definidas por

$$g(j) = I_{(x,\infty)}(j) \quad \text{e}$$

$$h(j) = \left[\frac{q_0(F, m)}{q_0(G, m)} \right]^j,$$

para todo $j \in \mathcal{N}^*$.

Pela caracterização de ordenação estocástica

$$F \leq_{estoc} G \Leftrightarrow \int_0^\infty f(x) dF(x) \geq \int_0^\infty f(x) dG(x),$$

para toda função não crescente f . Temos então que

$$\int_0^1 (1-x)^m dF(x) \geq \int_0^1 (1-x)^m dG(x) \quad \text{ou}$$

$$\frac{q_0(F, m)}{q_0(G, m)} = \frac{\int_0^1 (1-x)^m dF(x)}{\int_0^1 (1-x)^m dG(x)} \geq 1$$

o que implica que a função h é não decrescente. Como g é também não decrescente,

$$E[g(k)h(k)|G, m, w] \geq E[g(k)|G, m, w]E[h(k)|G, m, w]. \quad (5.4)$$

Desde que

$$E[g(k)h(k)|G, m, w] = \frac{C(F, m, w)}{C(G, m, w)} E[I_{(k>x)}|F, m, w], \quad (5.5)$$

$$E[g(k)|G, m, w] = E[I_{(k>x)}|G, m, w] \quad (5.6)$$

e

$$E[h(k)|G, m, w] = \frac{C(F, m, w)}{C(G, m, w)}, \quad (5.7)$$

por (5.4), (5.5), (5.6) e (5.7) provamos (i).

(ii) É igual a prova do Teorema 3.4.4, página 52. Dados F , m e w , sejam g e h funções reais definidas por

$$g(n) = I_{(x, \infty)}(n) \quad \text{e}$$

$$h(n) = (n - w)I_{[w, \infty)}(n),$$

para todo $n \in \mathcal{N}$. Temos que

$$\begin{aligned} & Cov[g(k), h(k)|F, m, w] \\ &= (w+1) \frac{C(F, m, w+1)}{C(F, m, w)} [E\{I_{(k>x)}|F, m, w+1\} - E[I_{(k>x)}|F, m, w]], \end{aligned}$$

que é positiva pois g e h são ambas funções não decrescentes.

(iii) Dados F, m, w e $x \in \mathfrak{R}$ sejam g e h funções reais definidas por

$$g(n) = \left[\frac{q_0(F, m+1)}{q_0(F, m)} \right]^n = \left[\frac{\int_0^1 (1-x)^{m+1} dF(x)}{\int_0^1 (1-x)^m dF(x)} \right]^n \quad \text{e}$$

$$h(n) = I_{(x, \infty)}(n),$$

para todo $n \in \mathcal{N}^*$. Desde que g é uma função decrescente e h é uma função não decrescente, segue que

$$E[g(k)|F, m, w]E[h(k)|F, m, w] \geq E[g(k)h(k)|F, m, w]. \quad (5.8)$$

Como

$$E[g(k)|F, m, w] = \frac{C(F, m+1, w)}{C(F, m, w)}$$

e

$$E[g(k)h(k)|F, m, w] = \frac{C(F, m+1, w)}{C(F, m, w)} E[I_{(k>x)}|F, m+1, w],$$

(5.8) implica que

$$E[I_{(k>x)}|F, m+1, w] \leq E[I_{(k>x)}|F, m, w].$$

◇

Como consequência, os momentos a posteriori de k seguem esta mesma regra de ordenação. Denotemos por $M_r(F, m, w)$ o r -ésimo momento da distribuição a posteriori de k .

Corolário 5.3.2 *Seja π uma distribuição a priori para k , própria ou imprópria uniforme em $\mathcal{N}^*$, e $r \in \mathcal{N}^*$.*

(i) *Para toda F e G funções de distribuição acumulada em $\mathfrak{R}$, $m \in \mathcal{N}^*$ e $w \in \mathcal{N}$, tal que $A_{\pi, w} \neq \emptyset$,*

$$F \leq_{\text{estoc}} G$$

implica que

$$M_r(F, m, w) \geq M_r(G, m, w);$$

(ii) para toda função de distribuição F em $\mathfrak{R}$, $m \in \mathcal{N}^*$ e $w \in \mathcal{N}$ tal que $A_{\pi, w+1} \neq \emptyset$,

$$M_r(F, m, w) \leq M_r(F, m, w+1);$$

(iii) para toda função de distribuição F em $\mathfrak{R}$, $m \in \mathcal{N}^*$ e $w \in \mathcal{N}$ tal que $A_{\pi, w} \neq \emptyset$,

$$M_r(F, m+1, w) \leq M_r(F, m, w).$$

Prova. A prova do item (i) é igual a prova do Corolário 3.4.2, página 50: $F \leq_{estoc} G$ implica que $k|F, m, w \geq_{estoc} k|G, m, w$ o que implica $f(k)|F, m, w \geq_{estoc} f(k)|G, m, w$ para toda função crescente f . Como $f(j) = j^r$, $j \in \mathcal{N}^*$ é crescente, provamos o resultado. As provas dos itens (ii) e (iii) são análogas. $\diamond$

O caso especial em que F é a função de distribuição Beta é natural no sentido dela ser conjugada com relação a uma análise de dados Bernoulli. Suponhamos que F seja Beta com parâmetros α' e α'' . Por (5.2) temos que a distribuição a posteriori de k é

$$\pi(k|\alpha', \alpha'', m, w) = \frac{\binom{k}{w} \left\{ \frac{\Gamma(\alpha' + \alpha'')}{\Gamma(\alpha' + \alpha'' + m)} / \frac{\Gamma(\alpha'')}{\Gamma(\alpha'' + m)} \right\}^k \pi(k)}{C(\alpha', \alpha'', m, w)} \quad (5.9)$$

onde

$$C(\alpha', \alpha'', m, w) = \sum_{j \geq w} \binom{j}{w} \left\{ \frac{\Gamma(\alpha' + \alpha'')}{\Gamma(\alpha' + \alpha'' + m)} / \frac{\Gamma(\alpha'')}{\Gamma(\alpha'' + m)} \right\}^j \pi(j). \quad (5.10)$$

Denotemos por $M_r(\alpha', \alpha'', m, w)$ seu r -ésimo momento. Para este caso particular temos a seguinte relação.

Corolário 5.3.3 *Seja π uma distribuição a priori para k , própria ou imprópria uniforme em $\mathcal{N}^*$, e $r \in \mathcal{N}^*$. Para todo $\alpha'_1 > 0$, $\alpha''_1 > 0$, $\alpha'_2 > 0$ e $\alpha''_2 > 0$, $m \in \mathcal{N}^*$ e $w \in \mathcal{N}$, tal que $A_{\pi,w} \neq \emptyset$.*

$$\alpha'_1 \leq \alpha'_2 \text{ e } \alpha''_1 \geq \alpha''_2$$

implicam que

$$P[k > x | \alpha'_1, \alpha''_1, m, w] \geq P[k > x | \alpha'_2, \alpha''_2, m, w] \quad -\infty < x < \infty,$$

e

$$M_r(\alpha'_1, \alpha''_1, m, w) \geq M_r(\alpha'_2, \alpha''_2, m, w).$$

Prova. Sejam F_1 uma função de distribuição Beta com parâmetros α'_1 e α''_1 , e F_2 uma função de distribuição Beta com parâmetros α'_2 e α''_2 . É fato que

$$\alpha'_1 \leq \alpha'_2 \text{ e } \alpha''_1 \geq \alpha''_2$$

implicam que

$$F_1 \leq_{estoc} F_2,$$

(veja Stoyan e Daley [61, Apêndice 1]). O teorema e o corolário anteriores completam a prova. $\diamond$

5.4 Superposição de um número desconhecido de modelos independentes

Nesta seção mostramos que posteriores explícitas como aquelas obtidas na Proposição 2.4.1, página 24, e na Proposição 2.4.3, página 26, bem como a monotonicidade crescente da distribuição a posteriori de k em função do número de espécies distintas, como no Teorema 3.4.4, página 52, são propriedades da estimação Bayesiana do número de componentes de uma superposição de modelos independentes e identicamente distribuídos. De fato. Consideremos a formulação geral de

modelos estatísticos para amostragem de espécies na seção 1.2. página 3. Se dados k e F , $X_1, \dots, X_k$ forem independentes e identicamente distribuídas com distribuição F , ou seja, para quaisquer $x_1, \dots, x_k$,

$$P[X_1 = x_1, \dots, X_k = x_k | k, F] = \prod_{i=1}^k P[X_i = x_i | F],$$

então para quaisquer $a_1, a_2, \dots \in \mathcal{N}$ tais que $\sum_{j \geq 1} a_j < \infty$,

$$\begin{aligned} & P[A_1 = a_1, A_2 = a_2, \dots | k, F] \\ &= \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} P[X_1 = x_1, \dots, X_k = x_k | k, F] \\ &= \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} \prod_{i=1}^k P[X_i = x_i | F] \\ &= \binom{k}{a_0, a_1, \dots} \prod_{r \geq 0} \{P[X_1 = r | F]\}^{a_r}, \end{aligned}$$

onde $a_0 = k - w$, $w = \sum_{r \geq 1} a_r$. Portanto, dados $a_1, a_2, \dots \in \mathcal{N}$ tais que $\sum_{r \geq 1} a_r < \infty$, a função de verossimilhança de k é tal que

$$L_{(a_1, a_2, \dots)}(k) \propto \binom{k}{w} \{P[X_1 = 0 | F]\}^k, \quad (5.11)$$

$k \geq w$. Observemos que (5.11) é uma forma generalizada de (2.6) na página 23, de (3.3) na página 45 e de (5.1) na página 106. Portanto, numa superposição de um número desconhecido k de modelos independentes e identicamente distribuídos, a distribuição a posteriori de k , dada uma amostra de espécies $a_1, a_2, \dots$, depende da amostra somente através de $w = \sum_{r \geq 1} a_r$ sendo atualizada da seguinte forma:

$$\begin{aligned} \pi(k | F, a_1, a_2, \dots) &= \pi(k | F, w) \\ &= \frac{\binom{k}{w} (P[X_1 = 0 | F])^k \pi(k)}{C(F, w)}, \end{aligned} \quad (5.12)$$

$k \geq w$, onde π é uma distribuição a priori para k , própria ou imprópria uniforme em $\mathcal{N}^*$, e

$$C(F, w) = \sum_{j \geq w} \binom{j}{w} (P[X_1 = 0|F])^j \pi(j). \quad (5.13)$$

Na seção 2.4, página 24, temos uma série de proposições relativas a estimação de k numa superposição de processos de Poisson. Substituindo $(\frac{\beta}{\beta+t})^\alpha$ por $P[X_1 = 0|F]$ nessas proposições todas elas continuam sendo válidas para qualquer superposição de modelos independentes e identicamente distribuídos com distribuição F qualquer.

Com relação as propriedades de ordenação, podemos afirmar que a distribuição a posteriori do número de modelos independentes, dada uma amostra de espécies, é estocasticamente crescente no número de espécies distintas observadas na amostra, independentemente da forma de $P[X_1 = 0|F]$. Isto também já está provado na demonstração do Teorema 3.4.4 na página 52; basta substituírmos $q_0(F, t)$ por $P[X_1 = 0|F]$.

Capítulo 6

Estimação Bayesiana do número de espécies de uma população finita

6.1 Introdução

Consideremos uma população finita de tamanho conhecido, N , com um número desconhecido de espécies, k . Vamos estudar a estimação Bayesiana de k com base nas frequências das frequências de uma amostra desta população. Vimos que há na literatura dois tipos de amostragem de espécies de uma população finita, a hipergeométrica na seção 1.3.4, página 8, e a Bernoulli-hipergeométrica na seção 1.3.5, página 9. Hill [35] considera a amostragem hipergeométrica, supõe que, dado k , as k frequências absolutas de cada classe da população, $N_1, \dots, N_k$, tem distribuição uniforme no conjunto $\{(f_1, f_2, \dots, f_k) : \sum_{i=1}^k f_i = N, f_i \geq 1, i = 1, 2, \dots, k\}$ e que k tem distribuição Binomial Negativa truncada em zero, obtendo a média e a variância a posteriori de k . Até o momento esta é a única referência de que temos conhecimento sobre abordagem Bayesiana do problema das espécies em população finita. Verificamos, neste capítulo, que nas formulações Bayesianas destes dois tipos de amostragem a distribuição a posteriori de k tem propriedades monotônicas, como as que temos obser-

vado nas formulações dos outros modelos. Na seção 6.2 consideramos uma amostra hipergeométrica de tamanho conhecido n e seguimos os passos da formulação Bayesiana convencional adotada nos capítulos anteriores, ou seja, supomos que as frequências populacionais têm distribuição Dirichlet-Multinomial simétrica com parâmetros N e α (uma conjugada com relação a amostras hipergeométricas). A estimação Bayesiana de k neste caso é equivalente aquela baseada numa amostra multinomial de espécies de tamanho n selecionada de uma população infinita, cujas proporções populacionais tem distribuição Dirichlet simétrica com parâmetro α . Na seção 6.3 consideramos a amostragem Bernoulli-hipergeométrica, mais especificamente, que os elementos da população fazem parte, ou não, da amostra independentemente e um total de n elementos da população são capturados, com as probabilidades de captura dos indivíduos variando segundo suas espécies. Supomos uma priori arbitrária para k e que, dado k , as probabilidades de captura são independentes das frequências populacionais das espécies, com as probabilidades tendo distribuição Beta com parâmetros α' e α'' e as frequências populacionais tendo distribuição Dirichlet-Multinomial simétrica com parâmetros N e α , onde $\alpha = \alpha' + \alpha''$. Com estas suposições a distribuição a posteriori de k depende de cinco parâmetros, α' , α'' , N , n e w . Mostramos que a distribuição a posteriori de k é decrescente em α' , decrescente em α , crescente em N se $\frac{n}{N} > \frac{\alpha'}{\alpha' + \alpha''}$, decrescente em N se $\frac{n}{N} < \frac{\alpha'}{\alpha' + \alpha''}$, decrescente em n e crescente em w .

6.2 Amostragem hipergeométrica

Dados k, N e $n \in \mathcal{N}^*$ tais que $N \geq n$ e $N \geq k$, e dado $(N_1, N_2, \dots, N_k)$ tal que $N_i \in \mathcal{N}$, $i = 1, 2, \dots, k$ e $\sum_{i=1}^k N_i = N$, seja $(X_1, X_2, \dots, X_k)$ um vetor aleatório k -dimensional com distribuição hipergeométrica

com parâmetros N, n e $(N_1, N_2, \dots, N_k)$. Então, dados k, N, n e $(N_1, \dots, N_k)$, a distribuição das frequências das frequências de uma amostra hipergeométrica é

$$P[A_1 = a_1, A_2 = a_2, \dots | k, N_1, N_2, \dots, N_k] \\ = \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots, k]} \prod_{i=1}^k \binom{N_i}{x_i} / \binom{N}{n}, \quad (6.1)$$

para quaisquer $a_1, a_2, \dots \in \mathcal{N}$ tais que $n = \sum_{r \geq 1} r a_r$ e $w = \sum_{r \geq 1} a_r \leq k$.

Suponhamos que k tenha uma distribuição de probabilidade qualquer π em $\{1, 2, \dots, N\}$ e que, dado k , $(N_1, \dots, N_k)$ tenha distribuição Dirichlet-Multinomial simétrica com parâmetros N e α . Sob estas suposições a distribuição marginal de $(X_1, \dots, X_k)$, dado k , é Dirichlet-Multinomial simétrica com parâmetros n e α e, portanto, dado k , a distribuição da amostra de espécies é uma soma de termos Dirichlet-multinomiais como em (4.11), página 77. Então, dado $A_1 = a_1, A_2 = a_2, \dots$, temos que a função de verossimilhança de k é proporcional a (4.12), página 78, isto é

$$L_{(a_1, a_2, \dots)}(k) \propto \binom{k}{w} \frac{\Gamma(k\alpha)}{\Gamma(k\alpha + n)},$$

$k = w, w + 1, \dots, N$. Logo, a distribuição a posteriori de k tem exatamente a mesma forma daquela que encontramos no capítulo 4 e, portanto, já sabemos como α, n e w interferem na estimação Bayesiana de k .

Observemos que, dado k , se $(N_1, \dots, N_k)$ tiver distribuição Dirichlet-Multinomial simétrica com parâmetros N e $\alpha = 1$, truncada a partir de $(1, \dots, 1)$, temos a modelagem de Hill[35], cuja priori é tal que $N_i \geq 1, i = 1, \dots, k$.

6.3 Amostragem Bernoulli-hipergeométrica

Consideremos uma população finita de tamanho conhecido, N , particionada em um número desconhecido de k classes $e_1, \dots, e_k$ de tamanhos desconhecidos $N_1, \dots, N_k$, respectivamente. Suponhamos que cada elemento da classe e_i faça parte de uma amostra selecionada da população com probabilidade desconhecida q_i , independentemente dos outros elementos da população, $i = 1, \dots, k$. Então, o tamanho da amostra é uma variável aleatória, onde a i -ésima classe e_i contribui com um número aleatório de elementos com distribuição Binomial (N_i, q_i) , $i = 1, \dots, k$. Logo, a distribuição da amostra de espécies é dada por: para quaisquer $a_1, a_2, \dots \in \mathcal{N}$ tais que $n = \sum_{r \geq 1} r a_r \in \{0, 1, \dots, N\}$ e $w = \sum_{r \geq 1} a_r \leq k$,

$$\begin{aligned} P[A_1 = a_1, A_2 = a_2, \dots | k, q_1, \dots, q_k, N_1, \dots, N_k] \\ = \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} \prod_{i=1}^k \binom{N_i}{x_i} q_i^{x_i} (1 - q_i)^{N_i - x_i}, \end{aligned} \quad (6.2)$$

onde o conjunto $[a_1, a_2, \dots; k]$ já foi definido nos capítulos anteriores.

O caso particular em que $q_1 = \dots = q_k = q$ já foi introduzido no capítulo 1, seção 1.3.5, página 9.

Consideremos agora a versão Bayesiana deste modelo, supondo que todas as quantidades desconhecidas $k, N_1, \dots, N_k, q_1, \dots, q_k$ sejam aleatórias. Suponhamos que π seja a distribuição a priori para k e que, dado k , $(N_1, \dots, N_k)$ seja independente de $(q_1, \dots, q_k)$ com $(N_1, \dots, N_k)$ tendo distribuição Dirichlet-Multinomial simétrica com parâmetros N e α e com $q_1, \dots, q_k$ independentes e identicamente distribuídas com distribuição Beta, com parâmetros $\alpha' > 0$ e $\alpha'' > 0$ tais que $\alpha' + \alpha'' = \alpha$. Então, a distribuição a priori de todos os parâmetros é dada por

$$\pi(N_1, \dots, N_k, q_1, \dots, q_k, k) = \pi(N_1, \dots, N_k, q_1, \dots, q_k | k) \pi(k) \quad (6.3)$$

$$= \frac{N! \Gamma(k\alpha)}{\Gamma(k\alpha + N)} \prod_{i=1}^k \frac{\Gamma(\alpha + N_i)}{N_i! \Gamma(\alpha)} \prod_{i=1}^k \frac{\Gamma(\alpha)}{\Gamma(\alpha') \Gamma(\alpha'')} q_i^{\alpha' - 1} (1 - q_i)^{\alpha'' - 1} \pi(k),$$

$N_i \in \mathcal{N}$, $i = 1, \dots, k$, $\sum_{i=1}^k N_i = N$, $q_i \in (0, 1]$, $i = 1, \dots, k$, $k \in \{1, 2, \dots, N\}$.

Sob as suposições (6.2) e (6.3), temos que

$$\begin{aligned} & P[A_1 = a_1, A_2 = a_2, \dots | k] \\ &= \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} E\left\{ \prod_{i=1}^k \binom{N_i}{x_i} E[q_i^{x_i} (1 - q_i)^{N_i - x_i} | N_1, \dots, N_k] \right\} \\ &= \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} E\left[\prod_{i=1}^k \binom{N_i}{x_i} \frac{\Gamma(\alpha' + x_i) \Gamma(\alpha'' + N_i - x_i)}{\Gamma(\alpha' + \alpha'' + N_i)} \frac{\Gamma(\alpha)}{\Gamma(\alpha') \Gamma(\alpha'')} \right] \\ &= \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} \sum_{N_1, \dots, N_k: \sum_{i=1}^k N_i = N} \prod_{i=1}^k \binom{N_i}{x_i} \frac{\Gamma(\alpha' + x_i) \Gamma(\alpha'' + N_i - x_i)}{\Gamma(\alpha' + \alpha'' + N_i)} \frac{\Gamma(\alpha)}{\Gamma(\alpha') \Gamma(\alpha'')} \frac{N! \Gamma(k\alpha)}{\Gamma(k\alpha + N)} \frac{\Gamma(\alpha + N_i)}{N_i! \Gamma(\alpha)} \\ &= \sum_{(x_1, \dots, x_k) \in [a_1, a_2, \dots; k]} \sum_{N_1, \dots, N_k: \sum_{i=1}^k N_i = N} \frac{(N - n)! \Gamma(k\alpha'')}{\Gamma(k\alpha'' + N - n)} \prod_{i=1}^k \frac{\Gamma(\alpha'' + N_i - x_i)}{(N_i - x_i)! \Gamma(\alpha'')} \prod_{i=1}^k \frac{\Gamma(\alpha' + x_i)}{\Gamma(\alpha') x_i!} \frac{N! \Gamma(k\alpha)}{\Gamma(k\alpha + N)} \frac{\Gamma(k\alpha'' + N - n)}{(N - n)! \Gamma(k\alpha'')} \\ &= \binom{k}{w} \binom{w}{a_1 a_2 \dots} \prod_{r \geq 1} \left[\frac{\Gamma(\alpha' + r)}{r! \Gamma(\alpha')} \right]^{a_r} \frac{N! \Gamma(k\alpha)}{\Gamma(k\alpha + N)} \frac{\Gamma(k\alpha'' + N - n)}{(N - n)! \Gamma(k\alpha'')}, \end{aligned}$$

para quaisquer $a_1, a_2, \dots \in \mathcal{N}$ tais que $n = \sum_{r \geq 1} r a_r \in \{0, 1, \dots, N\}$ e $w = \sum_{r \geq 1} a_r \leq k$.

Logo, para quaisquer $a_1, a_2, \dots \in \mathcal{N}$ tais que $w = \sum_{r \geq 1} a_r < \infty$, a função de verossimilhança de k é

$$L_{(a_1, a_2, \dots)}(k) \propto \binom{k}{w} \frac{\Gamma(k\alpha)}{\Gamma(k\alpha + N)} \frac{\Gamma(k\alpha'' + N - n)}{\Gamma(k\alpha'')},$$

$k \geq w$. Suponhamos que a distribuição a priori π para k seja uma função de probabilidade sobre $\{1, 2, \dots, N\}$ tal que $A_{\pi, w} \neq \emptyset$. Então a distribuição a posteriori de k , dados $a_1, a_2, \dots \in \mathcal{N}$ tais que $w = \sum_{r \geq 1} a_r < \infty$, é

$$\begin{aligned} & \pi(k|\alpha', \alpha'', N, n, w) \\ &= \binom{k}{w} \frac{\Gamma(k\alpha)}{\Gamma(k\alpha + N)} \frac{\Gamma(k\alpha'' + N - n)}{\Gamma(k\alpha'')} \pi(k) / C(\alpha', \alpha'', N, n, w), \end{aligned}$$

onde

$$\begin{aligned} & C(\alpha', \alpha'', N, n, w) \\ &= \sum_{j=w}^N \binom{j}{w} \frac{\Gamma(j\alpha)}{\Gamma(j\alpha + N)} \frac{\Gamma(j\alpha'' + N - n)}{\Gamma(j\alpha'')} \pi(j). \end{aligned}$$

Notemos que a posteriori de k depende de cinco parâmetros; mostraremos a seguir que ela é monótona em cada um destes parâmetros.

Teorema 6.3.1 *Sejam π uma distribuição a priori para k , $N \in \mathcal{N}^*$ e $n, w \in \mathcal{N}$ tais que $0 \leq w \leq n \leq N$ e $A_{\pi, w} \neq \emptyset$. Temos que,*

(i) fixado $\alpha > 0$, se α'_i e α''_i forem tais que $\alpha'_i + \alpha''_i = \alpha$, $i = 1, 2$, então

$$\alpha'_1 \leq \alpha'_2 \text{ ou equivalentemente } \alpha'_1/\alpha \leq \alpha'_2/\alpha$$

implica que

$$P[k > x|\alpha'_1, \alpha''_1, N, n, w] \geq P[k > x|\alpha'_2, \alpha''_2, N, n, w],$$

$$-\infty < x < \infty ;$$

(ii) fixado $\alpha'' > 0$, se $\alpha_1, \alpha_2, \alpha'_1$ e α'_2 forem tais que $\alpha'_i + \alpha'' = \alpha_i$, $i = 1, 2$, então

$$\alpha_1 \leq \alpha_2 \text{ ou equivalentemente } \alpha'_1/\alpha_1 \leq \alpha'_2/\alpha_2$$

implica que

$$P[k > x | \alpha'_1, \alpha'', N, n, w] \geq P[k > x | \alpha'_2, \alpha'', N, n, w],$$

$$-\infty < x < \infty.$$

Prova. (i) Fixados N, n, w e α , mostraremos que $\frac{\partial}{\partial \alpha''}(P[k > x | \alpha', \alpha'', N, n, w])$, $x \in \mathbb{R}$, é igual a covariância entre duas variáveis aleatórias positivamente correlacionadas. Seja $x \in \mathbb{R}$ e sejam F e G funções definidas por

$$F(\alpha', \alpha'', N, n, w) = \sum_{j=w}^N I_{(x, \infty)}(j) \binom{j}{w} \frac{\Gamma(j\alpha)}{\Gamma(j\alpha + N)} \frac{\Gamma(j\alpha'' + N - n)}{\Gamma(j\alpha'')} \pi(j) \quad (6.4)$$

e

$$G(\alpha', \alpha'', N, n, w) = \sum_{j=w}^N \binom{j}{w} \frac{\Gamma(j\alpha)}{\Gamma(j\alpha + N)} \frac{\Gamma(j\alpha'' + N - n)}{\Gamma(j\alpha'')} \pi(j), \quad (6.5)$$

para quaisquer $\alpha' > 0$, $\alpha'' > 0$ com $\alpha' + \alpha'' = \alpha$ e para quaisquer $N \in \mathcal{N}^*$ e $n, w \in \mathcal{N}$. Então,

$$P[k > x | \alpha', \alpha'', N, n, w] = \frac{F(\alpha', \alpha'', N, n, w)}{G(\alpha', \alpha'', N, n, w)}$$

e

$$\frac{\partial}{\partial \alpha''}(P[k > x | \alpha', \alpha'', N, n, w])$$

$$= \frac{\frac{\partial}{\partial \alpha''}(F(\alpha', \alpha'', N, n, w))}{G(\alpha', \alpha'', N, n, w)} - \frac{\frac{\partial}{\partial \alpha'}(G(\alpha', \alpha'', N, n, w))}{(G(\alpha', \alpha'', N, n, w))^2} F(\alpha', \alpha'', N, n, w).$$

Como pelo Lema 4.6.1, página 84, para todo $j = 1, 2, \dots, N$,

$$\frac{\partial}{\partial \alpha''} \left(\frac{\Gamma(j\alpha'' + N - n)}{\Gamma(j\alpha'')} \right) = \frac{\Gamma(j\alpha'' + N - n)}{\Gamma(j\alpha'')} \sum_{l=0}^{N-n-1} \frac{j}{j\alpha'' + l},$$

segue que

$$\begin{aligned} & \frac{\partial}{\partial \alpha''} (P[k > x | \alpha', \alpha'', N, n, w]) \\ &= Cov[I_{(k>x)}, \frac{\Gamma(k\alpha'' + N - n)}{\Gamma(k\alpha'')} \sum_{l=0}^{N-n-1} \frac{k}{k\alpha'' + l} | \alpha', \alpha'', N, n, w], \end{aligned}$$

que é positiva. Então, $\alpha_1'' \geq \alpha_2''$ implica que $P[k > x | \alpha_1', \alpha_1'', N, n, w] \geq P[k > x | \alpha_2', \alpha_2'', N, n, w]$, e como $\alpha_1' \leq \alpha_2'$ implica que $\alpha_1'' \geq \alpha_2''$, provamos (i). Vamos agora provar (ii). Fixados α'', N, n e w ,

$$\frac{\partial}{\partial \alpha'} (P[k > x | \alpha', \alpha'', N, n, w]) \leq 0,$$

pela prova do Teorema 4.6.2, página 84, e lembrando que $\alpha_i' + \alpha'' = \alpha_i$, $i = 1, 2$, segue o resultado. $\diamond$

Teorema 6.3.2 *Sejam π uma distribuição a priori para k , $\alpha' > 0$, $\alpha'' > 0$, tais que $\alpha' + \alpha'' = \alpha$, $N \in \mathcal{N}^*$ e $n, w \in \mathcal{N}$ tais que $0 \leq w \leq n \leq N$ e $A_{\pi, w} \neq \emptyset$. Então,*

$$(i) \text{ se } \frac{n}{N} > \frac{\alpha'}{\alpha' + \alpha''},$$

$$P[k > x | \alpha', \alpha'', N + 1, n, w] \geq P[k > x | \alpha', \alpha'', N, n, w] \text{ e}$$

$$(ii) \text{ se } \frac{n}{N} < \frac{\alpha'}{\alpha' + \alpha''},$$

$$P[k > x | \alpha', \alpha'', N + 1, n, w] \leq P[k > x | \alpha', \alpha'', N, n, w],$$

$$-\infty < x < \infty.$$

Prova. Dados α', α'' com $\alpha' + \alpha'' = \alpha$, N, n, w e $x \in \mathbb{R}$, sejam f e g funções reais definidas por

$$\begin{aligned} f(j) &= I_{(j, \infty)}(j) \text{ e} \\ g(j) &= \frac{j\alpha'' + N - n}{j\alpha + N}, \end{aligned}$$

$j \in \mathcal{N}$. Como

$$\frac{g(j+1)}{g(j)} > 1 \text{ se, e somente se, } \frac{n}{N} > \frac{\alpha'}{\alpha' + \alpha''},$$

para todo $j \in \mathcal{N}$, segue que

- (i) se $\frac{n}{N} > \frac{\alpha'}{\alpha' + \alpha''}$, a função g é crescente e
- (ii) se $\frac{n}{N} < \frac{\alpha'}{\alpha' + \alpha''}$, a função g é decrescente.

Logo, como a função f é não decrescente,

- (i) se $\frac{n}{N} > \frac{\alpha'}{\alpha' + \alpha''}$,

$$\text{Cov}[f(k), g(k) | \alpha', \alpha'', N, n, w] \geq 0 \text{ e}$$

- (ii) se $\frac{n}{N} < \frac{\alpha'}{\alpha' + \alpha''}$,

$$\text{Cov}[f(k), g(k) | \alpha', \alpha'', N, n, w] \leq 0.$$

Mas, pode-se verificar que

$$\begin{aligned} E[f(k)g(k) | \alpha', \alpha'', N, n, w] \\ = P[k > x | \alpha', \alpha'', N+1, n, w] \frac{C(\alpha', \alpha'', N+1, n, w)}{C(\alpha', \alpha'', N, n, w)}, \end{aligned}$$

$$\begin{aligned}
E[f(k)|\alpha', \alpha'', N, n, w] &= \\
&P[k > x|\alpha', \alpha'', N, n, w] \text{ e} \\
E[g(k)|\alpha', \alpha'', N, n, w] &= \frac{C(\alpha', \alpha'', N+1, n, w)}{C(\alpha', \alpha'', N, n, w)},
\end{aligned}$$

o que completa a prova. $\diamond$

Teorema 6.3.3 *Sejam π uma distribuição a priori para k , $\alpha' > 0$, $\alpha'' > 0$, tais que $\alpha' + \alpha'' = \alpha$, $N \in \mathcal{N}^*$ e $n, w \in \mathcal{N}$, tais que $0 \leq w \leq n < N$ e $A_{\pi, w} \neq \emptyset$. Então,*

$$P[k > x|\alpha', \alpha'', N, n+1, w] \leq P[k > x|\alpha', \alpha'', N, n, w].$$

Prova. Dados α', α'' , com $\alpha' + \alpha'' = \alpha$, N, n, w e $x \in \mathfrak{R}$, sejam f e g funções reais definidas por

$$\begin{aligned}
f(j) &= I_{(x, \infty)}(j) \text{ e} \\
g(j) &= \frac{1}{j\alpha'' + N - (n+1)},
\end{aligned}$$

$j \in \mathcal{N}$. A função f é não decrescente enquanto que a função g é decrescente, o que implica que

$$Cov[f(k), g(k)|\alpha', \alpha'', N, n, w] \leq 0.$$

Mas, pode-se verificar que

$$\begin{aligned}
E[f(k)g(k)|\alpha', \alpha'', N, n, w] &= \\
&= P[k > x|\alpha', \alpha'', N, n+1, w] \frac{C(\alpha', \alpha'', N, n+1, w)}{C(\alpha', \alpha'', N, n, w)},
\end{aligned}$$

$$\begin{aligned}
& E[f(k)|\alpha', \alpha'', N, n, w] \\
& = P[k > x|\alpha', \alpha'', N, n, w] \text{ e} \\
& E[g(k)|\alpha', \alpha'', N, n, w] \\
& = \frac{C(\alpha', \alpha'', N, n+1, w)}{C(\alpha', \alpha'', N, n, w)},
\end{aligned}$$

o que prova o resultado. $\diamond$

Teorema 6.3.4 *Sejam π uma distribuição a priori para k , $\alpha' > 0$, $\alpha'' > 0$ tais que $\alpha' + \alpha'' = \alpha$, $N \in \mathcal{N}^*$ e $n, w \in \mathcal{N}$ tais que $0 \leq w < n \leq N$ e $A_{\pi, w+1} \neq \emptyset$. Então,*

$$P[k > x|\alpha', \alpha'', N, n, w] \leq P[k > x|\alpha', \alpha'', N, n, w+1].$$

Prova. Dados α', α'' com $\alpha' + \alpha'' = \alpha$, N, n, w e $x \in \mathfrak{R}$, sejam f e g funções reais definidas por

$$\begin{aligned}
f(j) &= I_{(x, \infty)}(j) \text{ e} \\
g(j) &= I_{[w, \infty)}(j) \cdot (j - w),
\end{aligned}$$

$j \in \mathcal{N}$. Como f e g são funções crescentes,

$$\begin{aligned}
& Cov[f(k), g(k)|\alpha', \alpha'', N, n, w] \\
& = (w+1) \frac{C(\alpha', \alpha'', N, n, w+1)}{C(\alpha', \alpha'', N, n, w)} \{P[k > x|\alpha', \alpha'', N, n, w+1] - \\
& \quad P[k > x|\alpha', \alpha'', N, n, w]\} \geq 0
\end{aligned}$$

o que prova o resultado. $\diamond$

Comentários finais

Consideramos, nesta tese, formulações Bayesianas dos modelos estatísticos para amostragem de espécies mais frequentes na literatura. Mostramos que certas propriedades de monotonicidade da distribuição a posteriori do número de espécies, k , são válidas em quaisquer destas formulações bayesianas. Mostramos também que os estimadores dados pelos momentos a posteriori de k são consistentes. A validade destas propriedades em todos os modelos considerados é um forte indício de que elas devem ser válidas também para um número maior de modelos bayesianos para estimação do número de espécies, e para outras versões do problema das espécies, como a estimação do tamanho de uma população e até mesmo para versões bayesianas da estimação da dimensão de um espaço paramétrico.

Uma característica comum de todas as versões bayesianas que analisamos é que a distribuição a posteriori de k depende dos dados somente através do número de espécies distintas observadas na amostra, da opinião a priori sobre os parâmetros *nuisance* através de um hiperparâmetro, e de um parâmetro estabelecido pelo planejamento de amostragem. Isto não acontece se considerarmos prioris hierárquicas para os parâmetros *nuisance* pois, neste caso, a distribuição a posteriori de k depende globalmente das frequências das frequências, de uma dis-

tribuição a priori sobre o hiperparâmetro e do parâmetro estabelecido pelo planejamento de observação da amostra. Um atrativo dos modelos hierárquicos é que eles produzem estimativas para k que dependem das frequências das frequências levando em conta a heterogeneidade das frequências das espécies na amostra. Uma pergunta natural seria se os modelos hierárquicos também possuem propriedades de monotonicidade estocástica.

Uma outra questão que deve ser mencionada é o aspecto computacional das distribuições a posteriori de k em modelos bayesianos hierárquicos. Em geral o procedimento hierárquico nos proporciona posterioris para k sem uma forma explícita e com dificuldades não triviais nos cálculos de estimativas para k . Este tipo de dificuldade ocorre também na estimação bayesiana hierárquica do tamanho de uma população pelo método de captura e recaptura. Felizmente, George e Robert [26] mostram que a implementação de cálculos bayesianos no modelo de captura e recaptura é bem sucedida porque as distribuições condicionais necessárias para a implementação da amostragem Gibbs tem formas explícitas e, portanto, com amostras facilmente simuladas. Vimos nos capítulos 2, 3 e 5 que as posterioris obtidas para k dado um hiperparâmetro e dada uma amostra de espécies produzida por uma superposição de modelos independentes e identicamente distribuídos têm formas explícitas; estas distribuições com formas explícitas e conhecidas corresponderiam exatamente as distribuições condicionais que seriam usadas para gerar as amostras de Gibbs num cálculo Bayesiano hierárquico em dois estágios. Assim, acreditamos que a implementação da amostragem Gibbs para a determinação das quantidades relativas a posteriori de k , numa análise bayesiana hierárquica de uma superposição de k modelos independentes deverá ser também bem sucedida.

Referências Bibliográficas

- [1] ARNOLD, B. C. and BEAVER, R.J. (1988). Estimation of the number of classes in a population. *Biometrical Journal*, 30:413-424.
- [2] BOENDER, C. G. E. and RINNOOY KAN A. H. G. (1983). A Bayesian analysis of the number of cells of a multinomial distribution. *The Statistician*, 32:240-248.
- [3] BOENDER, C. G. E. and RINNOOY KAN A. H. G. (1987). A multinomial Bayesian approach to the estimation of population and vocabulary size. *Biometrika*, 74, 4: 849-856.
- [4] BUNGE, J. and FITZPATRICK, M. (1993). Estimating the number of species: a review. *J. Amer. Statist. Assoc.* 88: 364-373.
- [5] BURNHAM, K.P. and OVERTON, W.S. (1978). Estimation of the size of a closed population when capture probabilities vary among animals. *Biometrika*, 65:625-633.
- [6] CHAO, A. (1981). On estimating the probability of discovering a new species. *Annals of Statistics*, 9(6).
- [7] CHAO, A. (1984). Non-parametric estimation of the number of classes in a population. *Scand. J. Statist.*, 11: 265-270.
- [8] CHAO, A. (1987). Estimating the population size for capture-recapture data with unequal catchability. *Biometrics*, 43:783-791.

- [9] CHAO, A. and LEE, S. M. (1992). Estimating the number of classes via sample coverage. *J. Amer. Statist. Assoc.*, 87: 210-217.
- [10] CHAO, A., YIP, P. and HUEY-SHYAN, L. (1996). Estimating the number of species via a Martingale estimating function. *Statistica Sinica*, 6: 403-418.
- [11] CHAPMAN, D. G. (1951). Some properties of the hypergeometric distribution with applications to zoological sample censuses, in *University of California Publications in Statistics* (Vol.1), eds. M. Loeve, G. M. Kuznets, E. L. Lehmann, and J. Neyman, Berkeley and Los Angeles: University of California Press, pp. 131-160.
- [12] CHARALAMBIDES, C. A. and SINGH, J. (1988). A review of the Stirling number, their generalizations and statistical applications. *Communications in Statistics, Part A-Theory and Methods*, 17: 2533-2595.
- [13] CLAYTON, M. K. and FREES, E. W. (1987). Nonparametric Estimation of the probability of discovering a new species. *Journal of the American Statistical Association*, 82: 305-311.
- [14] COHEN, A., and SACKROWITZ, H. B. (1990). Admissibility of estimators of the probability of unobserved outcomes. *Annals of the Institute of Statistical Mathematics*, 42: 623-636.
- [15] CRAIG, C. C. (1953). Use of marked specimens in estimating populations. *Biometrika*, 40: 170-176.
- [16] DARROCH, J. N. (1958). The multiple-recapture census, I. Estimation of a closed population. *Biometrics*, 36: 149-153.
- [17] DRIML, M. and ULLRICH, M. (1967). Maximum likelihood estimate of the number of types. *Acta Technica Csav*, 3: 300-303.

- [18] EFRON.B. and THISTED.R. (1976). Estimating the number of unseen species: how many words did Shakeaspeare know? *.Biometrika*, 63: 435-447.
- [19] EMIGH, H. (1983). On the number of the observed classes from a multinomial distribution. *Biometrics*, 39: 485-491.
- [20] ENGEN,S. (1974). On species frequency models. *Biometrika*, 61: 263-270.
- [21] ENGEN, S. (1978). *Stochastic Abundance Models*, London:Chapman and Hall.
- [22] ESTY, W. W. (1985). Estimation of the number of classes in a population and the coverage of a sample. *Mathematical Scientist*, 10: 41-50.
- [23] ESTY, W. W. (1986). The size of a coinage. *Numismatic Cronicle*, 146: 185-215.
- [24] FRANK, O.(1978). Estimation of the number of connectede components in a graph by using a sampled subgraph. *Scandnavian Journal of Statistics, Theory and Applications*, 5: 177-188.
- [25] FISHER,R.A., CORBET,A.S. and WILLIANS,C.B. (1943). The relation between the number o species and the number of individuals in a random sample of an animal populatiom. *Journal Animal Ecology*, 12: 42-58.
- [26] GEORGE, E.I. and ROBERT, C.P. (1992). Capture-recapture estimation via Gibbs sampling. *Biometrika* 79: 677-683.
- [27] GOOD, I.J.(1950). *Probability and the Wheighting of Evidence*, London: Charles Griffin.

- [28] GOOD, I. J. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika* 40: 237-264.
- [29] GOOD, I. J. and TOULMIN, G. H. (1956). The number of new species and the increase in population coverage when a sample is increased. *Biometrika* 43: 45-63.
- [30] GOOD, I. J. (1965). *The Estimation Of Probabilities: An Essay On Modern Bayesian Methods*. Research Monograph No. 30, M.I.T. Press.
- [31] GOODMAN, L. A. (1949). On the estimation of the number of classes in a population. *Annals of Mathematical Statistics*, 20: 572-579.
- [32] GOODMAN, L. A. (1953). Sequential sampling tagging for population size problems. *Annals of Mathematical Statistics*, 24: 56-69.
- [33] HARRIS, B. (1968). Statistical Inference in the classical occupancy problem unbiased estimation of the classes. *J. Amer. Statist. Assoc.* 63: 837-847.
- [34] HARWIT, M. (1981). *Cosmic Discovery: The Search, Scope, and Heritage of Astronomy*, New York: Basic Books.
- [35] HILL, B.M. (1979). Posterior Moments of the number of species in a finite population and the posterior probability of finding a new species. *Journal of the American Statistical Association*, 74: 668-673.
- [36] HOLST, L. (1981). Some Asymptotic Results for incomplete multinomial or Poisson samples. *Scandinavian Journal of Statistics*, 8: 243-246.

- [37] IVCHENKO, G. J. and TIMONINA, E.E. (1982). Estimating the size of a finite population. *Theory of Probability and Its Applications*, 27: 403-406.
- [38] JOHNSON, N.L. and KOTZ, S. (1977). *Urn Models and Their Application*. New York: John Wiley.
- [39] JOAG-DEV, K. and PROSCHAN, F.(1983). Negative association of variables with applications. *The Annals of Statistics*, 11: 286-295.
- [40] KORWAR, R.M. (1988). On the Observed Number of classes from Multivariate Power Series and Hypergeometric Distribution. *Sankhya: The Indian Journal of Statistics*, 50: 39-59.
- [41] LEHMANN, E.L. (1966). Some concepts of dependence. *Annals of Mathematical Statistics*, 37: 1137-1153
- [42] LEITE, J. G. and BOLFARINE, H. (1994). Bayesian Estimation of the number of equiprobable classes in a population. *REBRAPE*, 8, 2, 127-133.
- [43] LEITE, J. G. and PEREIRA, C. A. B. (1990). Bayes estimation of the size of a finite population: capture/recapture sequential sample data. *International Statistical Review* 58, 3 : 201-213.
- [44] LEWINS, W.A. and JOANES, D.N. (1984). Bayesian estimation of the number of species. *Biometrics* 40: 323-328.
- [45] LEWONTIN, R.C. and PROUT, T. (1956). Estimation of the number of different classes in a population. *Biometrics* 12:211-223.
- [46] LISEK, B. (1978). Comparability of special distributions. *Math, Operat.-forschung Statist., Ser. Statist.* 9, 587-598.

- [47] LO.S. (1992). From the species problem to a general coverage problem via a new interpretation. *Annals of Statistics* 20:1094-1109.
- [48] MCNEIL, D. (1973). Estimating an author's vocabulary. *Journal of the American Statistical Association*, 68: 92-96.
- [49] MARSHAL, A. W. and OLKIN. I. (1979). *Inequalities: theory of majorization and its applications*. Academic Press, New York.
- [50] NAYAK, T.K. (1988). Estimating population size by recapture sampling. *Biometrika*, 75:113-120.
- [51] NAYAK, T.K. (1989). A note on estimating the number of errors in a system by recapture sampling. *Statistics and Probability Letters*, 7: 91-94.
- [52] NAYAK, T. K. (1991). Estimating the number of component processes of a superimposed process. *Biometrika* 78: 75-81.
- [53] NAYAK, T. K. and CHRISTMAN, M. C. (1992). Effect of unequal catchability on estimates of the number of classes in a population. *Scand. J. Statist.* 19: 281-287.
- [54] NAYAK, T.K. (1992). On statistical analysis of a sample from a population of unknown species. *Journal of Statistical Planning and Inference*, 31: 187-198.
- [55] NAYAK, T.K. (1993). On estimating the total rate of the unobserved processes. *Statistics and Probability Letters*, 17:351-354.
- [56] ROBBINS,H.E. (1968). Estimating the total probability of the unobserved outcomes of an experiment. *Annals of Mathematical Statistics*,39: 256-257.
- [57] ROSS, S.M. (1985a). Statistical estimation of software reliability. *IEEE Trans. Software Engineering*, SE-11 (5): 479-483.

- [58] ROSS, S.M. (1985b). Software reliability: the stopping rule problem. *IEEE Trans. Software Engineering*, SE-11 (12): 1472-1476.
- [59] STARR, N. (1979). Linear estimation of the probability of discovering a new species. *The Annals of Statistics*, 7: 644-652.
- [60] SHLOSSER, A. (1981). On estimation of the size of the dictionary of a long text on the basis of a sample. *Engineering Cybernetics*, 19: 97-102.
- [61] STOYAN, D. and DALEY, J. (1983). *Comparison Methods for Queues and Other Stochastic Models*. Wiley, N.Y.
- [62] TAYLOR, J.M. (1983). Comparisons of certain distribution functions. *Math. Operat.-forschung Statist., Ser. Statist.* 13.
- [63] ZIELINSKI, R. (1981). A statistical estimate of the structure of multi-extremal problems. *Mathematical Programming* 21: 348-356.
- [64] WANI, J.K. and LO, H.P. (1983a). A characterization of invariant power-series abundance distributions. *Canadian Journal of Statistics*, 11: 317-323.
- [65] WANI, J.K. and LO, H.P. (1983b). The structure of invariant abundance distributions. *Journal of the Indian Statistical Association*, 21: 1-7.
- [66] WONG, C. K. and YUE, P. C. (1973). A majorization theorem for the number of distinct outcomes in N independent trials. *Discrete Mathematics* 6: 391-398.
- [67] YIP, P. (1991). A method of inference for a capture-recapture experiment in discrete time with variable capture probabilities. *Communications in Statistics-Stochastic Models*, 7: 343-362.

- [68] YULE, G.U. (1944). *Statistical Study of Literary Vocabulary*.
Cambridge University Press.