

VANESSA KELLY DOS SANTOS

**UMA GENERALIZAÇÃO DA DISTRIBUIÇÃO DO ÍNDICE DE
DIVERSIDADE GENERALIZADO POR GOOD COM
APLICAÇÃO EM CIÊNCIAS AGRÁRIAS**

RECIFE-PE - MARÇO/2009



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

UMA GENERALIZAÇÃO DA DISTRIBUIÇÃO DO ÍNDICE DE DIVERSIDADE GENERALIZADO POR GOOD COM APLICAÇÃO EM CIÊNCIAS AGRÁRIAS

Dissertação apresentada ao Programa de Pós-Graduação em Biometria e Estatística Aplicada como exigência parcial à obtenção do título de Mestre.

Área de Concentração:

Modelagem Estatística e Computacional
(com ênfase nas áreas agrárias, biológicas e humanas)

Orientador: Prof. Dr. Rinaldo Luiz Caraciolo Ferreira

Co-orientador: Prof. Dr. José Antônio Aleixo da Silva

RECIFE-PE - MARÇO/2009.

Ficha catalográfica

S237g Santos, Vanessa Kelly dos
Uma generalização da distribuição do índice de diversidade generalizada por Good com aplicação em Ciências Agrárias/ Vanessa Kelly dos Santos. -- 2009.
57 f.

Orientador: Rinaldo Luiz Caraciolo Ferreira.
Monografia (Especialização em Biometria e Estatística Aplicada) – Universidade Federal Rural de Pernambuco.
Departamento de Estatística e Informática.
Inclui referências.

CDD 574.018

1. Diversidade
2. Índice de Shannon
3. Índice de Simpson
4. Generalização de Good
 - I. Ferreira, Rinaldo Luiz Caraciolo
 - II. Título

UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

**UMA GENERALIZAÇÃO DA DISTRIBUIÇÃO DO ÍNDICE DE DIVERSIDADE
GENERALIZADO POR GOOD COM APLICAÇÃO EM CIÊNCIAS AGRÁRIAS**

Vanessa Kelly dos Santos

Dissertação julgada adequada para obtenção do título de mestre em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 24/03/2009 pela Comissão Examinadora.

Co-orientador:

Prof. Dr. José Antônio Aleixo da Silva
Universidade Federal Rural de Pernambuco

Banca Examinadora:

Prof. Dr. Luiz Carlos Marangon
Universidade Federal Rural de Pernambuco

Prof. Dr. Eufrázio de Souza dos Santos
Universidade Federal Rural de Pernambuco

Prof. Dr. Tatijana Stosic
Universidade Federal Rural de Pernambuco

Dedico a minha tia Ivone, com carinho.

Agradecimentos

Agradeço profundamente a todas as pessoas que entraram na minha vida e me inspiraram, iluminaram e incentivaram, direta ou indiretamente, em mais esta conquista.

Também expresso minha gratidão às seguintes pessoas pelo enorme apoio e contribuição à minha jornada e criação desta dissertação.

À minha tia, Ivonete dos Santos, pelo exemplo de vida, pela luta em criar a sobrinha-filha e dar uma boa educação em meio a tantas pedras no meio do caminho e por ser a principal incentivadora dos meus estudos desde a minha opção pelo curso de Estatística.

Ao meu tio Ivo que sempre torceu e incentivou as minhas vitórias assim como a minha tia Jacira.

Aos meus irmãos Wimberto, Alessandro e Maria Eduarda por compreenderem o motivo da minha ausência, assim como minha mãe.

Aos amigos de longa data, em especial Milka e Silvia, presentes em minha defesa e Rafael e Maria Carolina, sempre presentes em minha vida apesar da distância física.

A todos os amigos conquistados na graduação, em especial: Alessandro, Erinaldo, Esdras, Felipe, Hemílio, pelo apoio incondicional, a Katiane, Maíra, Syntia, Tadeu, pela amizade e palavras de incentivo, nunca me deixando cair e, por fim, Wagner pela orientação em momentos de incerteza.

Principalmente, agradeço as amigas Amanda, que apesar de fazer parte de minha vida há tão pouco tempo foi e ainda será de vital importância para minha formação assim como Juliana com quem sempre pude contar.

À super mãe Zuleide pelo carinho e tratamento da alma e a Marco pela presteza.

Agradeço ainda as professoras Laélia e Tatijana pelas palavras de incentivo e orientação e ao Professor Gauss que me indicou a direção a seguir em um dos momentos críticos pelo qual tive que passar.

Ao meu orientador Rinaldo Caraciolo e ao meu co-orientador José Aleixo.

Ao Programa de Pós-graduação em Biometria e Estatística Aplicada e à capes pelo apoio financeiro.

E, por fim, Àquele sem o qual nada teria sido possível: Deus.

*“Somente quando for cortada a última árvore,
pescado o último peixe e poluído o último rio
é que as pessoas vão perceber que não se
pode comer dinheiro.”*

Provérbio indígena

Resumo

A diversidade como um conceito foi inicialmente introduzida por Williams em Fisher *et al.* (1943). Mais tarde, Good (1953, 1982) propôs um índice generalizado que incluía como casos especiais os índices de Shannon e o de Simpson. Baczkowski *et al.* (1997, 1998) propuseram generalizar essa generalização derivando os quatro primeiros momentos e obtendo assim uma distribuição para o índice antes generalizado por Good (BACZKOWSKI *et al.*, 2000). Sendo assim, apresenta-se uma nova generalização que, além de ter os índices de Shannon e Simpson como casos especiais, engloba índices mais gerais como o não familiar (PATIL & TAILIIE, 1982). Os momentos de $h(\alpha, \beta, \delta)$ aqui apresentados estendem os resultados apresentados em Baczkowski *et al.* e Bowman *et al.* para uma classe de índices de diversidade mais geral, concluindo-se então que enquanto a distribuição do índice de Shannon pode ser aproximado por uma distribuição Gaussiana, caso haja diferença entre abundância de espécies, para valores mais gerais de (α, β, δ) , sugere-se uma distribuição do tipo I como sendo a mais apropriada. Os resultados obtidos são tão consistentes quanto os de trabalhos que lidam com populações reais como em Heip & Engels (1974), principalmente quando examina-se o índice de Shannon.

Palavras-chave: Diversidade, índice de Shannon, índice de Simpson, generalização de Good.

Abstract

Diversity as a concept was first introduced by Williams in Fisher *et al.* (1943). Later, Good (1953, 1982) proposed a generalized index that included as special cases both Shannon's and Simpson's indices. Baczkowski *et al.* (1997, 1998) generalized the proposed generalization deriving the first four moments and then obtaining a distribution prior to the general index for Good (BACZKOWSKI *et al.*, 2000). Therefore, is proposed a new generalization that, in addition to the indices the Shannon and Simpson as special cases, includes more general indices such as the unfamiliar (PATIL & TAILIIE, 1982). The moments of $h(\alpha, \beta, \delta)$ presented here extend the results presented in Baczkowski *et al.* and Bowman *et al.* for a class of diversity indices of more general and it is concluded then that while the distribution of index Shannon can be approximated by a Gaussian distribution, if any difference between abundance of species, to more general values of (α, β, δ) , it is suggested a distribution of type I as the most appropriated. The results are also consistent with those presented for real populations, as in Heip & Engels (1974), especially when it examines the index of Shannon.

Keywords: Diversity, Shannon's index, Simpson's index, Good's generalization.

Sumário

1	INTRODUÇÃO	10
2	REVISÃO DE LITERATURA	12
2.1	Diversidade ecológica e suas medidas	12
2.2	Índices de diversidade	18
2.2.1	Índices de diversidade mais comumente utilizados	19
2.3	A construção e interpretação dos índices de diversidade de Shannon e Simpson via encontros inter-específicos	25
2.4	A Distribuição Multinomial	28
2.5	Assimetria e curtose	29
2.5.1	Assimetria	29
2.5.2	Curtose	30
2.6	A Distribuição Beta Tipo I	30
2.7	A Função Geratriz de Momentos	31
2.8	Expansões de Cornish-Fisher	32
3	MATERIAL E MÉTODO	33
3.1	A distribuição de um índice de diversidade generalizado por Good	33
3.2	Método para derivação dos momentos	35
4	RESULTADOS	37
4.1	Momentos não centrais para distribuição de abundância geral	39
4.2	Assimetria e curtose de $h(\alpha, \beta, \delta)$	46

5 APLICAÇÕES	49
6 CONSIDERAÇÕES FINAIS	51
REFERÊNCIAS BIBLIOGRÁFICAS	52

1 INTRODUÇÃO

O termo diversidade diz respeito à variedade de idéias, características ou elementos diferentes entre si, em determinado assunto, situação ou ambiente. Seu conceito é amplo, e existem aplicações em diferentes campos do conhecimento humano. Em ecologia, pode ser definido como o número de espécies que ocorrem em uma amostra tirada em uma unidade de área, volume de água, certo número de indivíduos, etc; em uma determinada unidade de tempo.

A Biodiversidade é o campo da biologia e meio ambiente que se refere ao estudo das relações quantitativas entre a riqueza de diferentes categorias biológicas e a abundância¹ relativa de espécies dentro das comunidades, incluindo variabilidade ao nível local (diversidade alfa), complementariedade biológica entre habitats (diversidade beta) e variabilidade entre paisagens (diversidade gama), agregando, desta forma, a totalidade dos recursos vivos e dos recursos genéticos, assim como seus componentes.

Os seres vivos de uma comunidade mantêm constantes relações entre si, exercendo assim influências recíprocas em suas vidas. As inter-relações podem ser evidenciadas entre indivíduos de uma mesma população, ou seja, entre indivíduos de uma mesma espécie, são as chamadas relações intra-específicas ou podem também ser evidenciadas entre populações diferentes, ou seja, entre indivíduos de espécies distintas, sendo, neste caso, conhecidas como relações interespecíficas. Em outras palavras, não existem vácuos ecológicos: organismos individuais nascem, crescem e eventualmente se reproduzem como membros de populações que integram comunidades biológicas dinâmicas. Os indivíduos que pertencem à uma mesma população costumam manter, entre si, relações ou interações intra-específicas, de intensa intimidade. Integrantes de uma mesma população local são por definição membros de uma mesma espécie, quer elas sejam amistosas (acasalamentos, relações entre mãe e filho etc.) ou não (disputas territoriais entre machos

¹É o número de indivíduos de uma determinada espécie existentes numa dada área num dado período de tempo, calculado em relação ao número total de indivíduos de todas as espécies existentes nessa área

adultos, brigas fraticidas entre filhotes etc.). Relações interespecíficas, entre indivíduos de espécies diferentes, costumam ser menos íntimas, mas ainda assim são de vital importância para a sobrevivência e o bem-estar dos envolvidos. Assim, é mesmo possível dizer que, para a maioria dos animais, interações interespecíficas são a chave para a sobrevivência, enquanto interações intra-específicas são a chave para o sucesso reprodutivo.

Tal diversidade é uma preocupação especial para biólogos, particularmente ecologistas e biogeógrafos. Vários autores já propuseram muitas medidas diferentes de diversidade, o que resultou em questões básicas ainda sem respostas.

Uma medida de diversidade é um parâmetro extremamente reducionista, que objetiva expressar toda a complexidade estrutural, a riqueza específica e as interações bióticas e abióticas de uma comunidade ecológica através de um único número, no entanto, ainda não há uma maneira bem estabelecida de expressar a diversidade de uma comunidade como um todo (MARTINS & SANTOS, 1999).

Os índices matemáticos podem ser classificados como a medida mais simples disponível para quantificar a diversidade de uma amostra, em que usam como base a riqueza de espécies. Em Loyd *et al.* (1968), Pielou considerou que dentre os vários índices, até então propostos, os mais comumente utilizados são os índices de Shannon-Wiener e Simpson que serão aqui trabalhados por apresentarem tendências similares na medição da diversidade.

Good (1953,1982) propôs um índice de diversidade generalizado que inclui como casos especiais os índices de Shannon e Simpson. Anos mais tarde, Baczowski *et al.* (1997, 1998) apresentaram uma generalização para tal índice generalizado, em que, a princípio, eles derivaram os quatro primeiros momentos do índice de diversidade, tanto para uma distribuição de abundância de espécies em geral, como para o caso em que todas as espécies tinham iguais abundâncias, o caso equiprovável, obtendo, em seguida, com base em suas medidas de assimetria e curtose, informações sobre a distribuição do índice. A partir deste, objetivou-se apresentar, nesta dissertação, uma generalização da distribuição de um índice de diversidade já generalizado por Good, através de possíveis inter-relações entre indivíduos, como apresentado em Patil & Taillie(1982).

2 REVISÃO DE LITERATURA

2.1 Diversidade ecológica e suas medidas

A diversidade como um conceito foi inicialmente introduzida por Williams em Fisher *et al.* (1943). Supondo que em todas, ou aproximadamente todas as coleções biológicas, a frequência de abundância de espécies tem uma distribuição log-series, ele propôs que o simples parâmetro desta distribuição, denotado por α , fosse tratado como uma medida de diversidade de uma coleção. Esse índice tem caído em desuso com a compreensão de que a distribuição de abundância de espécies pode ter muitas formas, e que um índice de distribuição livre é então desejável.

Uma medida de diversidade foi proposta por Khinchin (1957). Ele considerou uma população infinita de indivíduos que pode ser classificada em s classes, ou espécies, A_1, \dots, A_s . Todos os indivíduos pertencem a uma e somente uma classe, e a probabilidade que um indivíduo selecionado aleatoriamente pertença a classe A_j é p_j . Sendo assim, $\sum_{j=1}^s p_j = 1$. Como uma medida de diversidade de uma população, é desejável encontrar uma função de p_j , $H'(p_1, \dots, p_s)$, que satisfaça as seguintes condições:

1. Para um dado s a função assume seu maior valor quando $p_j = \frac{1}{s}$, $\forall j$. Denota-se este valor máximo por $L(S)$. Então,

$$L(S) = H' \left(\frac{1}{s}, \dots, \frac{1}{s} \right).$$

2. A diversidade da população é desconhecida se for postulada a existência de $(s+1)$ -ésima, $(s+2)$ -ésima, \dots , classes para as quais nenhum indivíduo pertence; isto é

$$H'(p_1, \dots, p_s, 0, \dots, 0) = H'(p_1, \dots, p_s).$$

3. Supondo que a população esteja sujeita a um processo de separação e classificação da população em t classes, B_1, \dots, B_t . Cada indivíduo pertence a exatamente uma das B classes e a probabilidade que este venha a pertencer a classe B_k é q_k , em que $\sum_{k=1}^t q_k = 1$. Então a dupla classificação apresenta st classes diferentes, $A_j B_k$, ($j = 1, \dots, s; k = 1, \dots, t$;) e a probabilidade de que um indivíduo escolhido aleatoriamente pertença a classe $A_j B_k$ pode ser descrita como π_{jk} . Claramente, se a classificação A e a classificação B são independentes, $\pi_{jk} = p_j q_k$, mas se os atributos cujas classificações são mutuamente dependentes então $\pi_{jk} = p_j q_{jk}$, em que, q_{jk} é a probabilidade condicional de um indivíduo pertencer a B_k dado que ele pertence a A_j . Para a diversidade da população duplamente classificada é possível escrever

$$H'(AB) = H'(\pi_{11}, \pi_{12}, \dots, \pi_{st}).$$

Em seguida, suponha $H'_j(B) = H'(q_{j1}, \dots, q_{jt})$ para a diversidade sobre a classificação B dentro da classe A_j e denote

$$H'_A(B) = \sum_{j=1}^s p_j H'_j(B),$$

como a diversidade média sobre a classificação B dentro de todas as classes A . Então, a condição 3 diz que deve haver

$$H'(AB) = H'(A) + H'_A(B). \quad (2.1)$$

Se as classificações são independentes, isto é, $q_{jk} = q_k \forall j$, tem-se

$$H'(AB) = H'(A) + H'(B). \quad (2.2)$$

Tendo especificado as três condições satisfatórias para H' , agora é possível notar que a única função com estas propriedades é

$$H'(p_1, \dots, p_s) = -\lambda \sum_{j=1}^s p_j \log(p_j), \quad (2.3)$$

em que λ é uma constante positiva.

Antes disso, porém, é desejável mostrar como tais condições são interpretadas em um contexto ecológico.

A condição 1 garante que para uma população com um dado número de espécies a medida de diversidade será máxima quando todas as espécies estiverem presentes em iguais proporções (ou com equidade¹ máxima).

A condição 2 assegura que dadas duas populações em que as espécies são representadas uniformemente, a população com um grande número de espécies terá alta diversidade.

Para entender as implicações da condição 3 imagine uma floresta contendo s distintas espécies de árvores e suponha que tais árvores são também classificadas de acordo com uma classe de alturas. A classificação A se refere a classificação das espécies e a B se refere a classificação das alturas. Deste modo é possível medir separadamente a diversidade de espécie das árvores, $H'(A)$, e sua diversidade de altura por classe, $H'(B)$. Também é possível calcular sua diversidade “espécie-altura”, $H'(AB)$, usando ambas classificações.

A condição 3 declara que se espécies e classe de alturas são inteiramente independentes como na Figura 2.1 (a) então $H'(AB) = H'(A) + H'(B)$. Neste caso, conhecer a espécie arbórea não dá informação sobre sua altura e vice-versa. Mas se árvores de uma espécie possuem alturas que pertencem a uma mesma classe, própria para a espécie (ou, equivalentemente, se árvores de uma dada classe de altura são todas de uma mesma espécie), como na Figura 2.1 (b), então $H'_B(A) = H'_A(B) = 0$ e a condição 3 agora requer que $H'(AB) = H'(A) = H'(B)$. Neste caso, conhecer a espécie arbórea revela imediatamente a classe de alturas a que está associada e vice-versa; nenhuma classificação adiciona informação ao que já foi dito anteriormente.

¹Equidade nada mais é que a homogeneidade entre abundância de espécies. Quanto maior a diversidade entre espécies maior a abundância e conseqüentemente maior a equidade

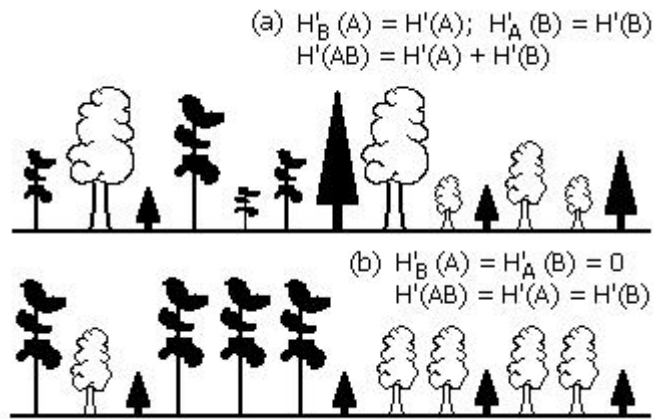


Figura 2.1: Ilustração de comunidades arbóreas

Partindo de (2.1) trabalhar-se-á com o caso especial em que $p_j = \frac{1}{s} \forall j$.

A partir das condições 1 e 2

$$L(s) = \left(\frac{1}{s}, \dots, \frac{1}{s} \right) = H' \left(\frac{1}{s}, \dots, \frac{1}{s}, 0 \right)$$

$$L(s) \leq H' \left(\frac{1}{s+1}, \dots, \frac{1}{s+1} \right)$$

$$L(s) \leq L(s+1),$$

mostrando que $L(s)$ é uma função não decrescente de s .

A seguir, considera-se m classificações da população, S_1, \dots, S_m , sendo estas mutuamente independentes, sobre cada uma das quais os indivíduos estão distribuídos em iguais proporções para r classes diferentes. Então, sobre a k -ésima classificação, S_k , por exemplo, a diversidade é $H'(S_k) = L(r)$. Tal relação é a mesma $\forall k$. Desde que as classificações sejam independentes, a diversidade da população sobre a classificação múltipla é dada por

$$H'(S_1 \dots S_m) = H'(S_1) + \dots + H'(S_m)$$

$$H'(S_1 \dots S_m) = mL(r)$$

Isto segue da condição 3, mas desde que a população tenha sido subdividida dentro de r classes iguais para cada m classificações sucessivas, a classificação múltipla resulta nas r^m últimas classes, cada uma contendo a mesma proporção de indivíduos. Portanto,

$$H'(S_1 \dots S_m) = L(r^m)$$

assim, nota-se que $mL(r) = L(r^m)$.

Igualmente, para algum outro par de inteiros positivos, n e s , tem-se

$$nL(s) = L(s^n).$$

Agora para valores arbitrários de r, s e n seja m escolhido de tal modo que $r^m \leq s^n \leq r^{m+1}$ ou $m \log(r) \leq n \log(s) \leq (m+1) \log(r)$.

Então,

$$\frac{m}{n} \leq \frac{\log(s)}{\log(r)} \leq \frac{m}{n} + \frac{1}{n}. \quad (2.4)$$

Relembrando que $L(s)$ é uma função não decrescente de s , como já foi mostrado, segue então, que

$$L(r^m) \leq L(s^n) \leq L(r^{m+1})$$

ou

$$mL(r) \leq nL(s) \leq (m+1)L(r),$$

em que

$$\frac{m}{n} \leq \frac{L(s)}{L(r)} \leq \frac{m}{n} + \frac{1}{n}. \quad (2.5)$$

De (2.4) e (2.5) segue agora que

$$\left| \frac{L(s)}{L(r)} - \frac{\log(s)}{\log(r)} \right| \leq \frac{1}{n}.$$

Portanto, desde que n possa ser tão grande quanto desejado

$$\frac{L(s)}{L(r)} = \frac{\log(s)}{\log(r)} \quad \text{ou} \quad L(s) \propto \log(s)$$

Então, ao escrever λ como uma constante de proporcionalidade têm-se

$$H'\left(\frac{1}{s}, \dots, \frac{1}{s}\right) = L(s) = \lambda \log(s) = \lambda \sum_{s=1}^n \left(\frac{1}{s}\right) \log\left(\frac{1}{s}\right),$$

como sendo a diversidade de uma espécie s da população em que todas as espécies são igualmente representadas.

Agora considera-se o caso geral em que os p_j 's não são todos iguais. Como descrito anteriormente, deseja-se definir a diversidade de uma população classificada dentro de s classes, A_1, \dots, A_s com uma proporção p_j de indivíduos na classe A_j . Seja $p_j = \frac{g_j}{g}$ em que g_j e g são inteiros. Em seguida, realiza-se uma segunda classificação dependente

da população, a classificação B , como segue: B tem g classes, todas de mesmo tamanho e classificação B_1, \dots, B_j . Estas classes são divididas em s grupos, com g_j classes no j -ésimo grupo. Se um indivíduo está na classe A_j sobre a classificação A é definido como pertencente ao j -ésimo grupo da classe B sobre a classificação B , já que $B \in A$. A probabilidade que ele pertença a alguma classificação dada nesse grupo é a mesma para todas as classes do grupo sendo igual a $\frac{1}{g}$ dada nesse grupo.

Agora nota-se que dentro da classe A_j a diversidade sobre a classificação B é

$$H'_j(B) = L(g_j) = \lambda \log(g_j).$$

Segue que $H'_A(B)$, a diversidade média sobre a classificação B dentro das classes A , é

$$H'_A(B) = \sum_{j=1}^s p_j H'_j(B) = \lambda \sum_{j=1}^s p_j \log(g_j).$$

$$H'_A(B) = \lambda \sum_{j=1}^s p_j \log(p_j) + \lambda \log(g).$$

desde que $g_j = g p_j$.

Agora considera-se a população duplamente classificada $A_j B_k$ para $j = 1, \dots, s$ e $k = 1, \dots, g$. O número total de classes com dupla classificação é $\sum_{j=1}^s g_j = g$ e estas classes são de igual tamanho. A diversidade de uma população sobre a dupla classificação é então

$$H'(AB) = L(g) = \lambda \log(g).$$

Agora, ao aplicar em (2.1) da condição 3, nota-se que

$$H'(A) = H'(AB) - H'_A(B)$$

$$H'(A) = \lambda \log(g) - \left[\lambda \sum_{j=1}^s p_j \log(p_j) + \lambda \log(g) \right]$$

$$H'(A) = -\lambda \sum_{j=1}^s p_j \log(p_j).$$

Sendo assim, fica provado que $-\lambda \sum_{j=1}^s p_j \log(p_j)$ está expressa apenas em função dos p_j 's que atendem as condições dadas anteriormente. Esta afirmação é verdadeira para valores racionais ou reais de p_j . A prova pode ser encontrada em Khinchin (1957).

Está fórmula foi proposta por Shannon (SHANNON & WEAVER, 1949) como uma medida de informação. No contexto ecológico H' mede a diversidade por indivíduos em uma população com muitas espécies, resta escolher as unidades em que a diversidade poderá

ser mensurada. O tamanho da unidade depende do valor atribuído para a constante λ e da base usada para os logaritmos.

2.2 Índices de diversidade

Um índice de diversidade é uma medida de “dispersão qualitativa” de uma população de indivíduos pertencentes a várias categorias qualitativamente diferentes (PIELOU, 1977). Do mesmo modo que estatísticas como a variância, o desvio-padrão, o desvio médio e a amplitude servem para medir a variabilidade quantitativa, índices de diversidade medem a variabilidade qualitativa (JOHNSON & KOTZ, 1988).

Os índices são classificados como medidas de riqueza (tipo I) - estimam o número de espécies s presentes em uma comunidade sem considerar sua estrutura de abundância; medidas de abundância (tipo II) - consideram a maneira como a abundância se distribui entre as espécies de uma comunidade já que tem por objetivo estimar o número de espécies e, ao mesmo tempo, manter as informações sobre sua abundância; e medidas de heterogeneidade (tipo III) - comparam a diversidade entre comunidades através de um modelo de distribuição de abundância (MARTINS & SANTOS, 1999).

Nem sempre todas as comunidades em comparação se ajustam a um mesmo modelo de distribuição de abundância. Quando isso acontece, então é mais conveniente usar uma medida de riqueza (índice tipo I) ou uma medida de heterogeneidade (índice tipo III). Segundo Martins & Santos (1999), o cálculo de um índice de heterogeneidade é muito mais simples e demanda menos tempo que o dos índices de riqueza ou de abundância.

Muitos autores acreditam que as medidas de abundância usam todas as informações levantadas de uma comunidade e representam a mais completa descrição matemática dos dados (MAGURRAN, 1988). No entanto, a maneira como a abundância se distribui entre as espécies é peculiar a cada comunidade, sendo necessário descrever tantas equações, quantas fossem as comunidades levantadas para, assim, descrever a distribuição de abundância entre espécies.

Com o intuito de facilitar a interpretação e o tratamento matemático dos dados foram considerados modelos de curvas de espécie-abundância cujo objetivo central é encontrar uma distribuição de probabilidade com um pequeno número de parâmetros que variam de uma comunidade para outra. Dentre os modelos de espécie-abundância, os mais citados na literatura são: “broken-stick” ou vara quebrada (MACARTHUR, 1957), lognormal (PRESTON, 1948), logsérie (FISHER *et al.*, 1943) e geométrico (WHITTAKER, 1967).

Na ecologia como um todo, inclusive no campo florestal, o modelo mais citado para estudar o comportamento da curva espécie-abundância é o de “broken-stick”, ou modelo da proporcionalidade de espaço (WEBB, 1974). Tal modelo foi proposto por MacArthur (1957) e baseia-se em interações competitivas entre as espécies de uma comunidade. Em comunidades constituídas por um pequeno número de espécies funcionalmente similares, em um ambiente relativamente homogêneo, o modelo de “broken-stick” tem dado bom ajuste aos dados coletados.

A abundância de cada espécie é determinada a partir de uma partição randômica dos recursos distribuídos ao longo de um contínuo de diferentes tipos de recursos. Os recursos são considerados como se estivessem distribuídos ao longo de um bastão em que cada segmento corresponde ao nicho² explorado por cada espécie ao longo do recurso considerado. O comprimento dos segmentos é proporcional à abundância relativa das espécies (PINTO-COELHO, 2006).

Sendo assim, a probabilidade da i -ésima espécie rara é aqui dada por,

$$p_i = \frac{1}{s} \sum_{r=1}^i \frac{1}{s-r+1} \quad (2.6)$$

em que $i = 1, \dots, s$.

Segundo Baczkowski (1997), no modelo de “broken-stick” a abundância é semelhante entre as espécies, ou seja, a equidade é grande. Quando tal abundância for igual para todas as espécies tem-se que o modelo é equiprovável, ou seja, a probabilidade de selecionar a i -ésima espécie em uma dada amostra é definida por $p_i = \frac{1}{s}$, para $i = 1, \dots, s$.

2.2.1 Índices de diversidade mais comumente utilizados

Números índices, que podem servir como medidas de diversidade, têm sido propostos. Os mais conhecidos são *Shannon-Wiener* e *Simpson*.

²Nicho é o conjunto de limites para todos os fatores ambientais importantes, dentro do qual a espécie consegue sobreviver e se reproduzir.

Índice de Shannon-Wiener (H')

O engenheiro elétrico e também matemático Claude Elwood Shannon (1916-2001), considerado o fundador da teoria da informação, publicou, no ano de 1948, o importante artigo científico intitulado *A Mathematical Theory of Communication* enfocando o problema de qual seria a melhor forma de codificar a informação que um emissor queira transmitir para um receptor. Neste artigo, ele trabalhou inclusive com as ferramentas teóricas utilizadas por Norbert Wiener, Claude Shannon propôs com sucesso uma medida de informação própria para medir incerteza sobre espaços desordenados (mais tarde complementada por Ronald Fisher).

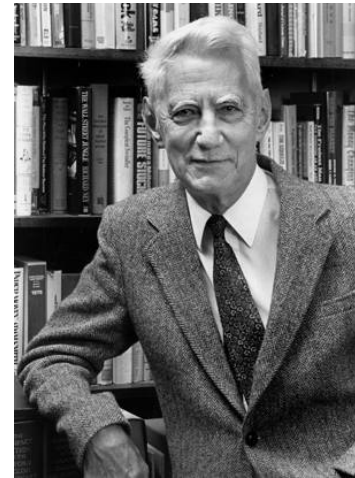


Figura 2.2: Claude Elwood Shannon.

Em 1949, Shannon, em co-autoria com o também matemático Warren Weaver (1894-1978), publicou o livro *The Mathematical Theory of Communication*, contendo reimpressões do seu artigo científico de 1948 de forma acessível também a não-especialistas, popularizando assim seus conceitos.

O índice de Shannon, também chamado de índice de Shannon-Weaver ou índice de Shannon-Wiener, é um dos vários índices de diversidade usados para medir a diversidade em dados categóricos. Trata-se da informação entrópica da distribuição, tratando espécies como símbolos e o tamanho da respectiva população como uma probabilidade. A vantagem desta medida de heterogeneidade é que ela leva em consideração o número das espécies e sua equitabilidade.

Seguindo a idéia de Simpson(1949), de que a medida de diversidade deve ser independente do modelo de distribuição de abundância e do tamanho da amostra, Margalef (1957, 1958) sugeriu o uso da teoria da informação (SHANNON & WEAVER, 1949) para estimar a diversidade de comunidades. Nessa teoria, a informação seria uma função do número de escolhas possíveis entre estados igualmente prováveis de certo nível organizacional, necessário para entendê-lo completamente: quanto maior o número de escolhas necessárias para entender a organização num certo nível, maior o conteúdo de informação.

Usando o índice de Shannon, H' , numa comunidade, a informação poderia ser baseada na atribuição de todos os indivíduos às suas respectivas espécies, por exemplo. Dessa maneira, a informação passaria a ser uma função do número de indivíduos. O conteúdo de informação H' dependeria não só do número total de indivíduos e do número total de espécies, mas também da proporção do número de indivíduos em cada espécie (p_j), embora p_j implique que as espécies não sejam igualmente prováveis (HAIRSTON, 1959), nem que o conteúdo de informação seja totalmente independente do tamanho da amostra (WHITTAKER, 1972).

Assim, supondo que a comunidade (tratada como indefinidamente grande) contenha s espécies e sendo p_j a proporção de indivíduos na j -ésima espécie ($j = 1, \dots, s$), o índice de Shannon-Wiener é definido por:

$$H' = - \sum_{j=1}^s p_j \log(p_j), \quad (2.7)$$

em que p_j é a abundância relativa de cada espécie, calculada através da razão entre o total de indivíduos de uma dada espécie j , n_j , e o número total de indivíduos na comunidade, N , ou seja: $p_j = \frac{n_j}{N}$. Vale ressaltar que $N = \sum_{j=1}^s n_j$.

O índice é derivado a partir da informação teórica. Dependendo da base logarítmica, H' é expresso em bits/indivíduo (base 2), nats/indivíduo (base natural) ou décits/indivíduo (base 10); o uso não segue um padrão específico. Quando se usa a base de logaritmos naturais, as propriedades matemáticas de H' apresentam muito mais consistência e coerência, de modo que há não só uma forte recomendação para usar nats/indivíduo (HUTCHESON, 1970), como também uma tendência mundial ao uso da base natural (MAGURRAN, 1988).

Pode ser mostrado que para um dado número de espécies, é possível obter um valor máximo para H' , $H'_{max} = \log(s)$, desde que todas as espécies estejam presentes em igual número. A prova de tal afirmação será apresentada a seguir.

Inicialmente, recomenda-se expandir o índice dado em (2.7). Assim sendo,

$$H' = - \sum_{j=1}^s \left(\frac{n_j}{N} \right) \log \left(\frac{n_j}{N} \right)$$

$$NH' = - \sum_{j=1}^s n_j [\log(n_j) - \log(N)]$$

$$NH' = - \sum_{j=1}^s n_j \log(n_j) + \log(N) \sum_{j=1}^s n_j$$

$$NH' - N \log(N) = - \sum_{j=1}^s n_j \log(n_j)$$

em seguida, desde que N seja uma constante positiva de um determinado tamanho populacional, e $N \log(N)$ seja uma constante, é possível definir

$$H_s = - \sum_{j=1}^s n_j \log(n_j), \quad (2.8)$$

em que a maximização de H_s equivale a maximização de H' .

Em suma, inicialmente particiona-se, arbitrariamente, uma população de um determinado tamanho em dois grupos, em que cada grupo recebe um número arbitrário de indivíduos e de espécies. Agora, dentro de cada grupo, cada espécie tem o mesmo número de indivíduos, mas o número dos indivíduos por espécies no primeiro grupo pode ser diferente do número de indivíduos por espécies no segundo grupo. Como foi provado que H_s é maximizada quando o número de indivíduos por espécies no primeiro grupo corresponde ao número de indivíduos por espécies no segundo grupo, está provado também que a população tem um índice máximo somente quando cada espécie na população é representada uniformemente, sendo assim satisfeitas as regras da indução matemática, H_s não depende da população total. Assim H_s pode ser facilmente construído ao somar os índices das duas subpopulações.

Índice de Simpson (D ou D')

O índice é derivado do “índice de concentração de Simpson” (SIMPSON, 1949), o qual supõe que dois indivíduos são independentemente retirados, ao acaso e sem reposição, de uma comunidade de s espécies contendo N indivíduos, em que N_j indivíduos pertencem a j -ésima espécie ($j = 1, \dots, s; \sum_{j=1}^s N_j = N$). Se a probabilidade de que ambos os indivíduos pertençam à mesma espécie for alta, é possível dizer que a comunidade apresenta um alto grau de concentração e tal probabilidade é utilizada como um índice de concentração, dito λ . Assim sendo, a probabilidade é dada por

$$\lambda = \sum_{j=1}^s \frac{N_j(N_j - 1)}{N(N - 1)}. \quad (2.9)$$

Para uma comunidade indefinidamente grande, o índice de Simpson é expresso como

$$D' = -\log \sum_{j=1}^s p_j^2 \approx -\log(\lambda'). \quad (2.10)$$

Para uma comunidade finita

$$D = -\log \left[\sum_{j=1}^s \frac{N_j(N_j-1)}{N(N-1)} \right] \equiv -\log(\lambda). \quad (2.11)$$

Dada uma comunidade inteiramente recenseada, λ é rigorosamente determinável sem erro amostral. Agora considera-se uma comunidade que de tão grande se torna infinita, dificultando assim sua representação e a estimação dos parâmetros através de uma amostra. Assume-se que uma amostra aleatória de comunidades individuais é avaliada embora, devido à renovação das espécies nas comunidades naturais, tal amostra possa ser difícil de ser obtida. Assume-se também que todas as espécies da comunidade são representadas na amostra, isto é, que $s^* = s$. Seja p_j a proporção, na comunidade original, de indivíduos da j -ésima espécie ($j = 1, \dots, s$). O valor real de p_j é desconhecido; seu estimador de máxima verossimilhança será $\hat{p}_j = N_j/N$. O valor real de λ é $\lambda = \sum_{j=1}^s p_j^2$ sendo sua estimação o objetivo principal em questão. Um estimador não tendencioso é dado por:

$$\tilde{\lambda} = \sum_{j=1}^s \frac{N_j(N_j-1)}{N(N-1)}, \quad (2.12)$$

como será apresentado a seguir (GOOD, 1953; HERDAN, 1958).

Ao denotar n_r o número de espécies representadas em uma amostra com r indivíduos; isto é, n_r é uma "freqüência da freqüência" e $\sum_r r n_r = N$. Em seguida, propõe-se $M = \sum_r r^2 n_r$; isto é, nota-se que $M = \sum_r N_j^2$. Assim, resta provar que $(M - N)/[N(N - 1)]$ é um estimador não viesado de $\sum_{j=1}^s p_j^2 = \lambda$. Inicialmente nota-se que

$$\frac{M - N}{N(N - 1)} = \frac{1}{N(N - 1)} \left[\sum_r r^2 n_r - \sum_r r n_r \right],$$

$$\frac{M - N}{N(N - 1)} = \frac{1}{N(N - 1)} \sum_r r(r - 1) n_r,$$

cujo valor esperado é dado por

$$E \left[\frac{M - N}{N(N - 1)} \right] = \frac{1}{N(N - 1)} \sum_r r(r - 1) E(n_r). \quad (2.13)$$

Agora, a cada momento um indivíduo é retirado da população. A probabilidade de que este

venha a pertencer a j -ésima espécie é p_j . Nesse caso, a probabilidade de que em uma amostra de tamanho N existam r representantes da j -ésima espécie será

$$P(r) = \binom{N}{r} p_j^r (1 - p_j)^{N-r}$$

Ao obter todas as s espécies estimadas na amostra, o valor esperado será representado por exatamente r indivíduos, isto é

$$E(n_r) = \sum_{j=1}^s \binom{N}{r} p_j^r (1 - p_j)^{N-r}.$$

Então, de (2.13),

$$E \left[\frac{M - N}{N(N - 1)} \right] = \sum_{j=1}^s \sum_{r=1}^R \frac{r(r-1)}{N(N-1)} \cdot \frac{N!}{r!(N-r)!} p_j^r (1 - p_j)^{N-r},$$

ou

$$E \left[\frac{M - N}{N(N - 1)} \right] = \sum_{j=1}^s p_j^2 \sum_{r=1}^R \binom{N-2}{r-2} p_j^{r-2} (1 - p_j)^{N-r}.$$

A nova soma da direita será 1 para todos os valores de j . Então,

$$E \left[\frac{M - N}{N(N - 1)} \right] = E \left[\frac{1}{N(N - 1)} \sum_{j=1}^s N_j(N_j - 1) \right] = \sum_{j=1}^s p_j^2 \quad (2.14)$$

Foi mostrado, então, que (2.12) provém de um estimador não viesado de λ , formalmente idêntico a (2.9), a fórmula do índice de concentração de toda uma comunidade. Vale recordar que λ mede a concentração, ou dominância, em uma comunidade e que adota-se esta medida como um índice de diversidade em que a melhor função de λ para esta proposta é $-\log(\lambda) = D$. A razão de usar D antes de qualquer outra função de λ que se intensifique com o seu decrescimento é que ambos D e a função de Shannon, H' , são casos especiais de uma função mais geral (PIELOU, 1975).

Na teoria da comunicação matemática, uma função geral (RENYI, 1961), H_α , chamada entropia de ordem α de um código composto de s diferentes tipos de símbolos e com os símbolos do j -ésimo tipo formando uma proporção p_j de todos os símbolos é definida por

$$H_\alpha = \frac{\log \left(\sum_{j=1}^s p_j^\alpha \right)}{1 - \alpha}.$$

Adotando $\alpha = 1$ e 2 encontra-se, respectivamente, que

$$H_1 = - \sum_{j=1}^s p_j \log(p_j) = H'$$

é o índice de Shannon-Wiener e

$$H_2 = - \log \left(\sum_{j=1}^s p_j^2 \right) = D'$$

é o índice de Simpson. Assim, o índice de Shannon e o índice de Simpson são intimamente relacionados.

O índice de Simpson não partilha com H' o mérito de fragmentar os grupos que já compõem a amostra à medida que se adicionam componentes e, portanto, não pode ser adaptado às medidas de diversidade hierárquica. Um dos mais interessantes tópicos de comunidades ecológicas comum é que se determina, para várias comunidades, qual nível de diversidade taxionômico³ é mais fortemente manifestado, e porque o alto nível de diversidade intra-táxon varia muito de uma comunidade para outra. Daí H' é, provavelmente, mais ecologicamente usado do que D . Porém, desde que a equidade se mostrou uma propriedade omissa e evasiva das comunidades talvez haja uma razão para que ambos H' e D resumam algumas das características quantitativas das comunidades com vasto número de espécies.

2.3 A construção e interpretação dos índices de diversidade de Shannon e Simpson via encontros inter-específicos

Ambos índices de diversidade ecológica atribuem zero de diversidade a uma comunidade de espécies únicas, sendo amplamente utilizados na literatura.

Pode-se interpretar os índices de Shannon e Simpson com base na noção de encontros interespecíficos. Ao reescrever o índice de diversidade de Simpson, D , na forma $\Delta_1 = \sum_{i=1}^s \pi_i(1 - \pi_i) = - \sum_{i=1}^s \pi_i R(\pi_i)$, também é possível interpretá-lo como a escassez média de espécies em uma comunidade, tendo a compreensão de que a escassez de espécies é medida por $R(\pi_i) = 1 - \pi_i$.

³Taxionomia é parte da botânica que classifica organismos vivos.

Wallace (1875) foi um dos primeiros a reconhecer a importância da diversidade na caracterização das comunidades e o primeiro a descobrir um membro da chamada i -ésima espécie. Ele encontrou em sua contínua jornada outros organismos, as vezes da mesma espécie e as vezes não. O provável evento mais raro entre os indivíduos da i -ésima espécie são os encontros interespecíficos. $R(\pi_i) = 1 - \pi_i$ é precisamente a probabilidade que um dado encontro é interespecífico. Patil & Taillie (1982) apresentam três diferentes cenários como sendo a exploração do conceito de encontro interespecífico versus intraespecífico.

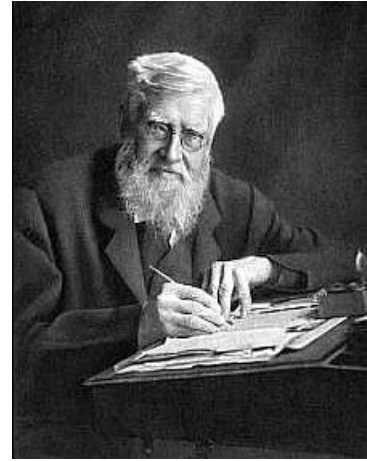


Figura 2.3: Alfred Russel Wallace.

Tempo de espera para um encontro intraespecífico: Com base nas descobertas de Wallace em suas viagens em busca da i -ésima espécie adota-se $Y + 1$ igual ao número de encontros até e inclusive o primeiro encontro intraespecífico. Sendo assim,

$$E[Y|\pi_i] = (1 - \pi_i)/\pi_i,$$

$$E[Y + 1|\pi_i] = 1/\pi_i,$$

$$E[1/(Y + 1)|\pi_i] = -\pi_i \log(\pi_i)/(1 - \pi_i),$$

quando

$$P(Y = y|\pi_i) = \pi_i(1 - \pi_i)^y, \quad y = 0, 1, \dots$$

Uma vez que Y seja grande, Y e $Y/(Y + 1)$ estão associadas a espécies raras sendo assim, medidas razoáveis de escassês. Tais medidas são variáveis aleatórias e deveriam ser substituídas por quantidades proporcionais. Existem vários caminhos para interpretar a “proporção” da razão; cada um implementando um índice diferente.

1. Índice de Simpson:

$$R(\pi_i) = E[Y|\pi_i]/E[Y + 1|\pi_i] = 1 - \pi_i,$$

$$\sum_{i=1}^s \pi_i R(\pi_i) = \sum_{i=1}^s \pi_i (1 - \pi_i). \quad (2.15)$$

2. Índice de Shannon:

$$R(\pi_i) = E[1/(Y + 1)|\pi_i] \cdot E[Y|\pi_i] = -\log(\pi_i),$$

$$\sum_{i=1}^s \pi_i R(\pi_i) = -\sum_{i=1}^s \pi_i \log(\pi_i). \quad (2.16)$$

3. Índice não familiar:

$$R(\pi_i) = E[Y/(Y + 1)|\pi_i] = 1 + \pi_i \log(\pi_i)/(1 - \pi_i),$$

$$\sum_{i=1}^s \pi_i R(\pi_i) = 1 + \sum_{i=1}^s \pi_i^2 \log(\pi_i)/(1 - \pi_i). \quad (2.17)$$

Tempo de espera para um encontro interespecífico: Agora, ainda com base nos conceitos apresentados por Wallace em meio a suas viagens quando em busca de uma nova espécie, define-se $Z + 1$ como sendo o número de encontros até e inclusive o primeiro encontro interespecífico. Assim sendo,

$$E[Z|\pi_i] = \pi_i/(1 - \pi_i),$$

$$E[Z + 1|\pi_i] = 1/(1 - \pi_i),$$

$$E[1/(Z + 1)|\pi_i] = -(1 - \pi_i) \log(\pi_i)/(1 - \pi_i).$$

Poucos Z 's estão associados com espécies raras e as variáveis de interesse são $1/Z$ e $1/(Z + 1)$.

1. Índice de Simpson:

$$R(\pi_i) = 1/E[Z + 1|\pi_i] = 1 - \pi_i.$$

Número fixado de encontros: Neste cenário, Y será o número de encontros inter-específicos e Z o número de encontros intraespecíficos que não um número N fixado de encontros.

1. Índice de Simpson:

$$R(\pi_i) = E[Y|\pi_i]/E[Y + Z|\pi_i] = 1 - \pi_i.$$

Algumas motivações biológicas para considerar encontros interespecíficos foram discutidas por Hulbert (1971), quem concluiu que se um índice tiver de ser usado, Simpson é conceitualmente mais apropriado que Shannon. Apesar de ser, no mínimo, curioso o índice de Shannon só aparece no primeiro cenário, todos os três índices, populares na literatura ecológica, podem dar uma interpretação comum sobre os encontros teóricos.

Vale ressaltar que apenas o primeiro cenário serviu de motivação para esta dissertação.

2.4 A Distribuição Multinomial

A distribuição Multinomial é considerada uma generalização da distribuição binomial e examina uma importante variável aleatória discreta com maior número de dimensões (MOOD, 1974). Refere-se a um experimento aleatório com s eventos mutuamente excluídos A_1, \dots, A_s que formam uma partição do espaço amostral, ou seja, quando o experimento for realizado garante-se que um, e somente um, dos eventos A_i ocorrerá. Sendo assim, denota-se $\pi_i = P(A_i)$ como a probabilidade do evento A_i ocorrer, supondo que π_i permaneça constante durante todas as repetições do experimento e garantindo assim que $\sum_{i=1}^s \pi_i = 1$.

Ao considerar n repetições independentes do experimento ε é possível representar o resultado dessas n provas por meio de um vetor aleatório s -dimensional (X_1, \dots, X_s) , em que a variável aleatória X_i corresponde ao número de vezes que o evento A_i ocorre nas n repetições de ε , sendo X_i variáveis aleatórias não independentes já que $\sum_{i=1}^s X_i = n$. Em consequência, logo que quaisquer $(s - 1)$ dessas variáveis aleatórias sejam conhecidas, o valor da outra ficará determinado, obtendo assim a seguinte definição:

Definição 1. Se X_1, \dots, X_s , forem definidas como acima, sua função de probabilidade é dada por

$$P(X_1 = n_1, \dots, X_s = n_s) = \frac{n!}{n_1! \dots n_s!} \pi_1^{n_1} \dots \pi_s^{n_s} \quad (2.18)$$

em que $\sum_{i=1}^s n_i = n$.

Propriedades: Supondo que (X_1, \dots, X_s) tenha uma distribuição multinomial, dada pela equação (2.18). Nesse caso

$$E(X_i) = n\pi_i$$

$$Var(X_i) = n\pi_i(1 - \pi_i) \text{ e}$$

$$Cov(X_i, X_j) = -n\pi_i\pi_j$$

Em termos ecológicos, considera-se um experimento que pode resultar em somente um número finito s de distintas possíveis espécies identificadas em uma determinada área em uma dada seleção.

2.5 Assimetria e curtose

As medidas de assimetria e curtose são frequentemente usadas para descrever o padrão característico de uma distribuição (JOANES & GILL, 1998), dentre outras coisas. A família de distribuições de Pearson (PEARSON & MERRINGTON, 1951) foi caracterizada pelos quatro primeiros momentos, sendo a assimetria e a curtose medidas apropriadamente usadas para selecionar uma amostra significativa dessa família.

2.5.1 Assimetria

Na teoria da probabilidade e Estatística, assimetria é o grau de desvio, ou afastamento da simetria, da distribuição de probabilidade de uma variável aleatória. Se a cauda esquerda apresentar mais informações que a cauda direita, é dito que a função é assimétrica à esquerda, ou negativa. Do mesmo modo, quando a cauda direita estiver mais pesada será chamada de assimétrica a direita ou positiva.

A assimetria pode ser definida através do terceiro momento padronizado, ou seja,

$$\alpha_3 = \mu_3 / \sigma^3 \tag{2.19}$$

em que μ_3 é o terceiro momento sobre a média e σ é o desvio padrão. Ou equivalentemente, pode ser definida como a razão da terceira parte da raiz quadrada do segundo cumulante, ou seja,

$$\sqrt{\beta_1} = \frac{k_3}{k^{3/2}}. \tag{2.20}$$

2.5.2 Curtose

É uma medida de dispersão que caracteriza o achatamento de uma curva da distribuição de probabilidade de uma variável aleatória. É definida como uma forma normalizada do quarto momento central μ_4 de uma distribuição. O quarto momento padronizado é definido como

$$\alpha_4 = \mu_4 / \sigma^4 \quad (2.21)$$

em que μ_4 é o quarto momento sobre a média e σ é o desvio padrão. Em trabalhos mais antigos ainda é possível ver a curtose definida como em (2.21), no entanto, atualmente é mais comum defini-la em função de cumulantes. Neste caso, a curtose corresponde ao quociente entre o quarto e o quadrado do segundo cumulante, ou seja,

$$\beta_2 = \frac{k_4}{k_2^2} = \frac{\mu_4}{\sigma^4} - 3. \quad (2.22)$$

que é conhecida como curtose de excesso. A constante que aparece ao término da equação (2.22) é explicada, frequentemente, como uma correção igualando a zero a curtose da distribuição normal.

Vale salientar que os cumulantes de uma distribuição são obtidos ao aplicar o logaritmo à função geratriz de momentos da variável aleatória de interesse expandindo-a em t , ou seja, $K = \log M_X(t)$.

2.6 A Distribuição Beta Tipo I

A família de distribuições de Pearson (PEARSON & MERRINGTON, 1951) é caracterizada pelos quatro primeiros momentos. Os coeficientes de assimetria e curtose podem ser usados para selecionar um membro apropriado para esta família de distribuições como o modelo de distribuições de $H(\alpha, \beta, \delta)$.

A distribuição beta do tipo I tem função densidade de probabilidade dada como segue

$$f(x) = \frac{x^{p-1}(1-x)^{q-1}}{B(p, q)}, \quad 0 \leq x \leq 1; \quad p, q > 0$$

em que $B(p, q)$ é a função beta usual. É possível estimar h de uma distribuição do tipo I, com classe (a_1, a_2) , transformando-a em uma distribuição beta do tipo I usando a relação

$$x = \frac{a_1 + h}{a_1 + a_2},$$

em que $a_2(p - 1) = a_1(q - 1)$, sendo possível escrever $c = a_1 + a_2$.

2.7 A Função Geratriz de Momentos

Seja $r = 1, 2, \dots$, o momento de ordem r de uma variável aleatória X é dado por $E(X^r)$, desde que o valor esperado exista. Se $E(X) = \mu < \infty$ é possível definir o momento de ordem central r por $E[(X - \mu)^r]$, sempre que essa quantidade exista.

Uma função que representa todos os momentos da variável aleatória X é conhecida como função geratriz de momentos e sua definição é dada a seguir:

Definição 2. Seja X uma variável aleatória cujo valor esperado e variância são conhecidos. Sendo assim, é possível definir a função geratriz de momentos $M_X(t)$ através da equação

$$M_X(t) = E(e^{tX}) \quad (2.23)$$

desde que a esperança seja finita para um t real em algum intervalo $-t_0 < t < t_0$ em que $t_0 > 0$. Na verdade, caso a fgm exista $\forall -t_0 < t < t_0, t_0 > 0$ é possível afirmar que $E(X^r)$ existe para $r = 1, 2, \dots$, ou seja,

$$E(X^r) = \frac{\partial^r}{\partial t^r} M_X(t) \Big|_{t=0} \quad (2.24)$$

Definição 3. Sejam X_i, X_j variáveis aleatórias definidas em um mesmo espaço de probabilidade e $t_i, t_j \in \mathfrak{R}$. A função geratriz de momentos bivariada, ou bidimensional, de tais variáveis é dada por

$$M_{X_i X_j}(t_i, t_j) = E(e^{t_i X_i + t_j X_j}) \quad (2.25)$$

desde que a esperança seja finita para os t_j 's tomados numa vizinhança de zero. Sendo assim, é possível obter os momentos das variáveis aleatórias X_i, X_j como segue

$$E(X_i^l X_j^m) = \frac{\partial^{l+m}}{\partial t_i^l \partial t_j^m} M_{X_i X_j}(t_i, t_j) \Big|_{t=0}. \quad (2.26)$$

2.8 Expansões de Cornish-Fisher

As Expansões de Cornish-Fisher são usadas para determinar numericamente as distribuições de probabilidade de estatísticas quando suas distribuições exatas são difíceis de ser computadas (CORDEIRO, 1999). Supondo que uma variável aleatória contínua padronizada Y tem média zero e variância um é possível supor também que y_α e u_α são definidas por $P(Y \leq y_\alpha) = \Phi(u_\alpha) = 1 - \alpha$.

As duas expansões assintóticas de Cornish-Fisher relacionam os quantis y_α e u_α , sendo uma a expansão normalizadora que expressa u_α como função de y_α e a outra sua expansão inversa dando y_α em termos de u_α . Sendo assim, é possível expressar u_α em função de y_α até $O(n^{-1})$ como segue

$$u_\alpha = p(y_\alpha) = y_\alpha - \frac{\rho_3}{\sqrt{n}}(y_\alpha^2 - 1) + \frac{\rho_3^2}{36n}(4y_\alpha^2 - 7y_\alpha) - \frac{\rho_4}{24n}(y_\alpha^3 - 3y_\alpha). \quad (2.27)$$

O polinômio (2.27) de Cornish-Fisher representa a transformação normalizadora $p(Y)$ da variável Y até $O(n^{-1})$, isto é, $p(Y) \sim N(0, 1) + (O_p^{-3/2})$, usado comumente para normalizar qualquer distribuição de probabilidade fazendo algum dos seus parâmetros tender para infinito, ou seja, substituindo-se n no resultado assintótico de grandes amostras ($n \rightarrow \infty$) por algum parâmetro da distribuição de interesse que cresce indefinidamente.

O objetivo da expansão inversa de Cornish-Fisher na inferência é expressar os quantis y_α de Y como função dos correspondentes quantis u_α da distribuição normal reduzida. A inversão da expansão (2.27) para calcular y_α em termos do quantil u_α da normal reduzida é feita através da fórmula geral de inversão de Lagrange (CORDEIRO, 1999). Então, $y_\alpha = u_\alpha + g(y_\alpha)$ pode ser expandido em termos de u_α como

$$y_\alpha - u_\alpha = g(u_\alpha) + \frac{Dg^2(u_\alpha)}{2!} + \frac{D^2g^3(u_\alpha)}{3!} + \dots \quad (2.28)$$

Identificando o polinômio $g(y_\alpha) = y_\alpha - p(y_\alpha)$ em (2.27), substituindo em (2.28) e calculando as potências de $g(u_\alpha)$ e suas derivadas, obtém-se y_α em função de u_α até $O(n^{-1})$ como

$$y_\alpha = u_\alpha + \frac{\rho_3}{\sqrt{n}}(y_\alpha^2 - 1) - \frac{\rho_3^2}{36n}(2u_\alpha^3 - 5u_\alpha) + \frac{\rho_4}{24n}(u_\alpha^3 - 3u_\alpha) \quad (2.29)$$

3 MATERIAL E MÉTODO

Os dados utilizados nesta dissertação foram tirados de artigos publicados em revistas brasileiras. A primeira base de dados foi coletada no Parque Estadual Mata dos Godoy (PEMG) localizado no município de Londrina, PR, sul do Brasil, ver Bianchini *et al.* (2003). Utilizou-se o método de parcelas em que o critério mínimo de inclusão dos indivíduos amostrados foi de 15cm de perímetro do tronco à altura do peito (PAP). Os indivíduos amostrados foram identificados com plaquetas metálicas numeradas, sendo registrados a altura total, estimada visualmente por comparação com a extensão telescópica de uma tesoura de poda alta de 7,5m, e o perímetro (PAP), medido com auxílio de fita métrica.

A segunda base de dados compõe um estudo desenvolvido em ambiente de terra firme, no parque fenológico do Campo Experimental da Embrapa Amazônia Ocidental, localizado na BR-174, Km 54, no Distrito Agropecuário da Suframa - DAS, Manaus-AM, o qual possui aproximadamente 400 ha. Os indivíduos foram avaliados segundo sua distribuição diamétrica, utilizando-se o intervalo de classe de 10cm como mostra Silva *et al.* (2008).

3.1 A distribuição de um índice de diversidade generalizado por Good

Good (1953, 1982) propôs um índice de diversidade generalizado que inclui como casos especiais tanto o índice de Shannon quanto o de Simpson.

Supondo que uma população consiste de s espécies com abundâncias relativas $\pi = (\pi_1, \dots, \pi_s)$ Good (1953, 1982) sugeriu medir a diversidade usando um índice na forma

$$H(\alpha, \beta) = \sum_{i=1}^s (\pi_i)^\alpha \{-\log(\pi_i)\}^\beta, \quad (3.1)$$

definida para valores inteiros não negativos de α e β , vale ressaltar que em termos nota-

cionais $-\log = -\ln$.

A equação (3.1) objetiva apresentar uma medida de diversidade mais geral que inclua, como casos especiais, tanto o índice de Shannon (1948) ou entropia, $H(1,1)$, quanto o índice de Simpson (1949), $H(2,0)$, abrangendo ainda a classe de índices relatados por Hill (1973), quando $\beta = 0$, que podem ser chamados de momentos constantes da população. Em outras palavras, $H(1,n)$ é o momento sobre zero.

Em termos práticos, uma amostra de tamanho n é avaliada, em que n_i corresponde ao número de indivíduos pertencentes a espécie i . A abundância relativa da i -ésima espécie, π_i , pode ser estimada através da estimativa de máxima verossimilhança, $p_i = n_i/n$, e o índice de diversidade generalizado será estimado por

$$h(\alpha, \beta) = \sum_{i=1}^s (p_i)^\alpha \{-\log(p_i)\}^\beta. \quad (3.2)$$

Baczkowski *et al.* (1997,1998) implementaram a generalização do índice de Good tal que (α, β) podem assumir valores no plano real \mathfrak{R}^2 . Eles determinaram uma série de valores (α, β) para cada $H(\alpha, \beta)$ que satisfizessem duas propriedades chave propostas por Pielou (1975), são elas,

P_1 : para s fixado, o índice aumenta a medida que as abundâncias relativas tornam-se mais uniformes;

P_2 : se as abundâncias relativas são uniformes então o índice é uma função crescente de s .

A região de valores válidos, ou apenas região aceitável, para (α, β) é apresentada em Baczkowski *et al.* (1997,1998), em que $H(\alpha, \beta)$ está contido na região $\alpha \in (0, 1]$ com β variando entre zero e um limite superior. Para uma inversa de $H(\alpha, \beta)$ a região é dita aceitável se satisfizer $\alpha \geq 1$ e $\beta \leq 0$, salientando-se que o caso em que (α, β) for igual a $(0, 1)$ não é considerado.

Na prática, ter conhecimento sobre as mudanças que ocorrem no índice de diversidade ao mover-se de um habitat para outro, é mais importante do que ter conhecimento prévio da diversidade de uma particular comunidade, em um tempo específico. Em Baczkowski *et al.* (1997,1998) foi sugerido que escolher diferentes valores para α e β acarreta particulares variações perceptíveis na diversidade. Tal variação pode ocorrer na riqueza ¹ ou na abundância relativa de espécies.

¹Defini-se riqueza de espécies como a abundância numérica em uma determinada comunidade.

Segundo Baczkowski *et al.* (2000), para testar se uma variação observada no índice de diversidade amostrado é significativa, faz-se necessário obter informações sobre a média e variância de $h(\alpha, \beta)$. O objetivo principal desse artigo foi avaliar a média $\mu = E[h(\alpha, \beta)]$ e os momentos centrais $\mu_r(h) = E[h(\alpha, \beta) - \mu]^r$ para $r = 2, 3, 4$, sendo esses usados para determinar a assimetria e curtose de $h(\alpha, \beta)$.

3.2 Método para derivação dos momentos

Os momentos de $h(\alpha, \beta)$ podem ser usados como uma ferramenta que sugere satisfatórias distribuições aproximadas para índice de diversidade, como mostram Baczkowski *et al.* (2000).

Ao derivar momentos, inicialmente, faz-se necessário descrever a função com a qual se trabalha, neste caso, $g(\pi; \alpha, \beta) = \pi^\alpha \{-\log(\pi)\}^\beta$, tal que, $h(\alpha, \beta) = \sum_{i=1}^s g(p_i)$. Em seguida, com base em Basharin (1959), expande-se $g(p; \alpha, \beta)$ em série de Taylor sobre π , estando esta agora em função da abundância amostral de espécies. Sendo assim, a expansão em série de Taylor retorna

$$g(p; \alpha, \beta) \approx g(\pi; \alpha, \beta) + \sum_{u=1}^{\infty} (p - \pi)^u g^{(u)}(\pi; \alpha, \beta) / u! \quad (3.3)$$

em que $g^{(u)}(\pi; \alpha, \beta)$ denota a u -ésima derivada de $g(\pi; \alpha, \beta)$ com respeito a π .

Agora escreve-se $\phi_i = p_i - \pi_i$, segundo Bowman *et al.* (1971), os momentos de ϕ_i são determinados ao notar que (n_1, \dots, n_s) modelam uma distribuição multinomial com parâmetros n e π_1, \dots, π_s .

Define-se $b = h(\alpha, \beta) - H(\alpha, \beta)$, tal que $E[b]$ é o viés de $h(\alpha, \beta)$. Segue que tanto μ quanto $\mu_r(h)$ podem então ser escritos em termos de $E[b^r]$, desde que

$$b = \sum_{i=1}^s \sum_{u=1}^{\infty} \phi_i^u g^{(u)}(\pi_i; \alpha, \beta) / u!. \quad (3.4)$$

Assim sendo,

$$E(b) = E \left[\sum_{i=1}^s \sum_{u=1}^{\infty} \phi_i^u g^{(u)}(\pi_i; \alpha, \beta) / u! \right]. \quad (3.5)$$

Como $\phi_i = p_i - \pi_i$ e $p_i = \frac{n_i}{n}$ então $\phi_i = \frac{1}{n}(n_i - n\pi_i)$. Portanto,

$$E(b) = \sum_{i=1}^s \sum_{u=1}^{\infty} \frac{1}{n^u} E (n_i - n\pi_i)^u \frac{g^{(u)}(\pi_i; \alpha, \beta)}{u!}. \quad (3.6)$$

Para obter os momentos não centrais de ordem r do viés, faz-se necessário avaliar termos de até $O(n^{-2})$, obtendo, conseqüentemente, a variância, assimetria e curtose de $h(\alpha, \beta)$. Feito isso é possível selecionar um membro apropriado para compor a família de distribuições de Pearson e apresentar uma distribuição para o índice de diversidade generalizado, como mostra a próxima seção, resultados.

4 RESULTADOS

Ao aplicar os métodos apresentados na seção anterior é possível apresentar, com base na generalização feita por Baczkowski *et al.* (2000), uma nova generalização que, além de ter os índices de Shannon e Simpson como casos especiais, engloba índices mais gerais como o não familiar apresentado em Patil e Taillie (1982).

Desta forma, seguindo a idéia de Good (1953,1982), como mostrada na equação 3.1, apresenta-se, nesta dissertação, uma forma generalizada para tal expressão resultando na equação

$$H(\alpha, \beta, \delta) = \sum_{i=1}^s \frac{(\pi_i)^\alpha \{-\log(\pi_i)\}^\beta}{(1 - \pi_i)^\delta} \quad (4.1)$$

em que π_i refere-se a abundância relativa das espécies na comunidade. Consequentemente, a estimativa deste índice é dada por

$$h(\alpha, \beta, \delta) = \sum_{i=1}^s \frac{(p_i)^\alpha \{-\log(p_i)\}^\beta}{(1 - p_i)^\delta} \quad (4.2)$$

em que p_i refere-se a abundância relativa das espécies na amostra.

Para obter os momentos centrais, ou seja, $E(n_i - n\pi_i)^u$ para $u = 2, 3, 4$, faz-se necessário, inicialmente expandir a equação (3.3) em série de Taylor, sobre π , sendo, neste caso, $g(\pi; \alpha, \beta) = \pi^\alpha \{-\log(\pi)\}^\beta / (1 - \pi)^\delta$, tal que, $h(\alpha, \beta, \delta) = \sum_{i=1}^s g(p_i)$ obtendo

$$g(p; \alpha, \beta, \delta) \approx g(\pi; \alpha, \beta, \delta) + \sum_{u=1}^{\infty} (p - \pi)^u g^{(u)}(\pi; \alpha, \beta, \delta) / u! \quad (4.3)$$

em que $g^{(u)}(\pi; \alpha, \beta, \delta)$ denota a u -ésima derivada de $g(\pi; \alpha, \beta, \delta)$ com respeito a π .

Ao escrever $\phi_i = p_i - \pi_i$ é possível determinar os momentos de ϕ_i sabendo-se que (n_1, \dots, n_s) modelam uma distribuição multinomial com parâmetros n e π_1, \dots, π_s , desde que $b = h(\alpha, \beta, \delta) - H(\alpha, \beta, \delta)$ possa ser escrito como

$$b = \sum_{i=1}^s \sum_{u=1}^{\infty} \phi_i^u g^{(u)}(\pi_i; \alpha, \beta, \delta) / u!. \quad (4.4)$$

O viés de $h(\alpha, \beta, \delta)$, representado por $E[b]$, permite que tanto μ quanto $\mu_r(h)$ possam ser escritos em termos de $E[b^r]$, assim sendo,

$$E(b) = \sum_{i=1}^s \sum_{u=1}^{\infty} \frac{1}{n^u} E(n_i - n\pi_i)^u \frac{g^{(u)}(\pi_i; \alpha, \beta, \delta)}{u!}. \quad (4.5)$$

em que

$$E(n_i - n\pi_i)^2 = n\pi_i(1 - \pi_i)$$

$$E(n_i - n\pi_i)^3 = n\pi_i(2\pi_i^2 - 3\pi_i + 1)$$

$$E(n_i - n\pi_i)^4 = n\pi_i - 7n\pi_i^2 + 3n^2\pi_i^2 + 12n\pi_i^3 - 6n^2\pi_i^3 - 6n\pi_i^4 + 3n^2\pi_i^4.$$

Após a obtenção dos momentos centrais para $r = 2, 3, 4$, é possível expandir a equação (4.5), sendo essa dada por

$$\begin{aligned} E(b) = & \sum_{i=1}^s \frac{1}{n^2} [\pi_i(1 - \pi_i)] \frac{g^{(2)}(\pi_i; \alpha, \beta, \delta)}{2!} + \sum_{i=1}^s \frac{1}{n^3} [\pi_i(2\pi_i^2 - 3\pi_i + 1)] \frac{g^{(3)}(\pi_i; \alpha, \beta, \delta)}{3!} \\ & + \sum_{i=1}^s \frac{1}{n^4} [n\pi_i - 7n\pi_i^2 + 3n^2\pi_i^2 + 12n\pi_i^3 - 6n^2\pi_i^3 - 6n\pi_i^4 + 3n^2\pi_i^4] \frac{g^{(4)}(\pi_i; \alpha, \beta, \delta)}{4!}. \end{aligned} \quad (4.6)$$

Ao ignorar os termos superiores a $O(n^{-2})$, a equação final do viés é representada por

$$\begin{aligned} E(b) = & \sum_{i=1}^s \pi_i(1 - \pi_i) \frac{g^{(2)}(\pi_i; \alpha, \beta, \delta)}{2n} + \sum_{i=1}^s \pi_i(2\pi_i^2 - 3\pi_i + 1) \frac{g^{(3)}(\pi_i; \alpha, \beta, \delta)}{6n^2} \\ & + \sum_{i=1}^s \pi_i(\pi_i^2 - 2\pi_i + 1) \frac{g^{(4)}(\pi_i; \alpha, \beta, \delta)}{8n^2}. \end{aligned} \quad (4.7)$$

em que $g^{(2)}(\pi; \alpha, \beta, \delta)$, $g^{(3)}(\pi; \alpha, \beta, \delta)$ e $g^{(4)}(\pi; \alpha, \beta, \delta)$ denotam a segunda, terceira e quarta derivadas em função de π_i , respectivamente.

4.1 Momentos não centrais para distribuição de abundância geral

Tendo conhecimento sobre a expressão do viés, dada por (4.7), é possível obter os momentos não centrais para $r = 1, 2, 3, 4$, para o caso em que as abundâncias entre espécies diferem. Assim,

$$\begin{aligned}
E(b) = & -\frac{1}{2} \{ \alpha(\alpha-1)[H(\alpha, \beta, \delta) - H(\alpha-1, \beta, \delta)] + \beta(\beta-1)[H(\alpha, \beta-2, \delta) - H(\alpha-1, \beta-2, \delta)] \\
& - (2\alpha-1)\beta[H(\alpha, \beta-1, \delta) - H(\alpha-1, \beta-1, \delta)] - 2\beta\delta[H(\alpha+1, \beta-1, \delta+1) - H(\alpha, \beta-1, \delta+1)] \\
& + 2\alpha\delta[H(\alpha+1, \beta, \delta+1) - H(\alpha, \beta, \delta+1)] + \delta(\delta+1)[H(\alpha+2, \beta, \delta+2) - H(\alpha+1, \beta, \delta+2)] \} n^{-1} \\
& + \{ \beta(\beta-1)(\beta-2)(\beta-3)H(\alpha-2, \beta-4, \delta)/8 + \beta(\beta-1)(\beta-2)(7-6\alpha)H(\alpha-2, \beta-3, \delta)/12 \\
& + \beta(\beta-1)(6\alpha^2 - 14\alpha + 7)H(\alpha-2, \beta-2, \delta)/8 - \beta(6\alpha^3 - 21\alpha^2 + 21\alpha - 5)H(\alpha-2, \beta-1, \delta)/12 \\
& + \alpha(\alpha-1)(\alpha-2)(3\alpha-5)H(\alpha-2, \beta, \delta)/24 - \beta(\beta-1)(\beta-2)(\beta-3)H(\alpha-1, \beta-4, \delta)/4 \\
& - \beta(\beta-1)(\beta-2)\delta H(\alpha-1, \beta-3, \delta+1)/4 + (\alpha-1)\beta(\beta-1)(\beta-2)H(\alpha-1, \beta-3, \delta) \\
& + (3\alpha-4)\beta(\beta-1)\delta H(\alpha-1, \beta-2, \delta+1)/2 + \beta(\beta-1)(-6\alpha^2 + 12\alpha - 5)H(\alpha-1, \beta-2, \delta)/4 \\
& - \beta\delta(9\alpha^2 - 21\alpha + 10)H(\alpha-1, \beta-1, \delta+1)/6 + \beta(\alpha-1)(2\alpha^2 - 4\alpha + 1)H(\alpha-1, \beta-1, \delta)/2 \\
& + \alpha(\alpha-1)^2\delta H(\alpha-1, \beta, \delta+1)/2 - \alpha(\alpha-1)^2(\alpha-2)H(\alpha-1, \beta, \delta)/4 \\
& + \beta(\beta-1)(\beta-2)(\beta-3)H(\alpha, \beta-4, \delta)/8 + \beta(\beta-1)(\beta-2)\delta H(\alpha, \beta-3, \delta+1) \\
& + \beta(\beta-1)(\beta-2)(5-6\alpha)H(\alpha, \beta-3, \delta)/12 + 3\beta(\beta-1)(\delta+1)\delta H(\alpha, \beta-2, \delta+2)/4 \\
& - 3(2\alpha+1)\beta(\beta-1)\delta H(\alpha, \beta-2, \delta+1)/2 + \beta(\beta-1)(6\alpha^2 - 10\alpha + 3)H(\alpha, \beta-2, \delta)/8 \\
& + \beta(\delta+1)\delta H(\alpha, \beta-1, \delta+2)/4 + \beta\delta(3\alpha^2 + 12\alpha - 7)H(\alpha, \beta-1, \delta+1)/6 \\
& + \beta(-6\alpha^3 + 15\alpha^2 - 9\alpha + 1)H(\alpha, \beta-1, \delta)/12 + \alpha(3\alpha-1)(\delta+1)\delta H(\alpha, \beta, \delta+2)/4 \\
& - \alpha(\alpha-1)(2\alpha-1)\delta H(\alpha, \beta, \delta+2)/2 + \alpha(\alpha-1)(\alpha-2)(3\alpha-1)H(\alpha, \beta, \delta)/24 \\
& - 3\beta(\beta-1)(\delta+1)\delta H(\alpha+1, \beta-2, \delta+2)/2 + \beta(\beta-1)\delta(9\alpha-4)H(\alpha+1, \beta-2, \delta+1)/8 \\
& - \beta(\delta+2)(\delta+1)\delta H(\alpha+1, \beta-1, \delta+3)/2 + 3\alpha\beta(\delta+1)\delta H(\alpha+1, \beta-1, \delta+2)/2 \\
& - \alpha\beta\delta(3\alpha-2)H(\alpha+1, \beta-1, \delta+1)/2 + (3\alpha+1)(\delta+2)(\delta+1)\delta H(\alpha+1, \beta, \delta+3)/6 \\
& - 3\alpha(\alpha+1)(\delta+1)\delta H(\alpha+1, \beta, \delta+2)/4 + \alpha^2(\alpha-1)\delta H(\alpha+1, \beta, \delta+1)/2 \\
& + 3\beta(\beta-1)(\delta+1)\delta H(\alpha+2, \beta-2, \delta+2)/4 + \beta(\delta+2)(\delta+1)\delta H(\alpha+2, \beta-1, \delta+3) \\
& - \beta(\delta+1)\delta(6\alpha+1)H(\alpha+2, \beta-1, \delta+2)/4 + (\delta+3)(\delta+2)(\delta+1)\delta H(\alpha+2, \beta, \delta+4)/8 \\
& - (\delta+2)(\delta+1)\delta(2\alpha+1)H(\alpha+2, \beta, \delta+3)/2 + \alpha(\delta+1)\delta(3\alpha+1)H(\alpha+2, \beta, \delta+2)/4 \\
& - \beta(\delta+2)(\delta+1)\delta H(\alpha+3, \beta-1, \delta+3)/2 - (\delta+3)(\delta+2)(\delta+1)\delta H(\alpha+3, \beta, \delta+4)/4 \\
& + (\delta+2)(\delta+1)\delta(3\alpha+2)H(\alpha+3, \beta, \delta+3)/6 + (\delta+3)(\delta+2)(\delta+1)\delta H(\alpha+4, \beta, \delta+4)/8 \} n^{-2}
\end{aligned}$$

$$\begin{aligned}
E(b^2) = & 4\alpha\delta(-\delta + \alpha(4 + 9\delta))H(2\alpha, 2\beta, 2\delta + 2) + 4\beta\delta(-1 - 4\delta + \beta(4 + 9\delta))H(2\alpha, 2\beta - 2, 2\delta + 2) \\
& - 4\beta\delta(-\delta + 2\alpha(\beta(4 + 9\delta)))H(2\alpha, 2\beta - 1, 2\delta + 2) - 16\alpha^2(2\alpha - 1)\delta H(2\alpha, 2\beta, 2\delta + 1) \\
& + 8\beta(4\beta^2 - 5\beta + 1)\delta H(2\alpha, 2\beta - 3, 2\delta + 1) - 4\beta(1 - 4\beta + 2\alpha(12\beta - 5))\delta H(2\alpha, 2\beta - 2, 2\delta + 1) \\
& + 32\alpha(3\alpha - 1)\beta\delta H(2\alpha, 2\beta - 1, 2\delta + 1) + 6(\alpha - 1)^2\alpha^2 H(2\alpha - 2, 2\beta, 2\delta) \\
& + 2\beta^2(3\beta^2 - 8\beta + 5)H(2\alpha - 2, 2\beta - 4, 2\delta) - 4\beta(\beta - 1)(-3\beta + \alpha(6\beta - 2))H(2\alpha - 2, 2\beta - 3, 2\delta) \\
& + 2\beta(\alpha(6 - 18\beta) + 3\beta + 2\alpha(9\beta - 4))H(2\alpha - 2, 2\beta - 2, 2\delta) - 12\alpha(2\alpha^2 - 3\alpha + 1)\beta H(2\alpha - 2, 2\beta - 1, 2\delta) \\
& + 4\alpha(2\alpha - 1)(3\alpha - 1)\delta H(2\alpha - 1, 2\beta, 2\delta + 1) - 8\beta(3\beta^2 - 4\beta + 1)\delta H(2\alpha - 1, 2\beta - 3, 2\delta + 1) \\
& + 4\beta(2 - 5\beta + 2\alpha(9\beta - 4))\delta H(2\alpha - 1, 2\beta - 2, 2\delta + 1) - 4(18\alpha^2 - 10\alpha + 1)\beta\delta H(2\alpha - 1, 2\beta - 1, 2\delta + 1) \\
& - 4\alpha^2(\alpha - 1)(2\alpha - 1)H(2\alpha - 1, 2\beta, 2\delta) - 4\beta^2(\beta - 1)(2\beta - 3)H(2\alpha - 1, 2\beta - 4, 2\delta) \\
& + 4\beta(\beta - 1)(8\alpha\beta - 2\alpha - 3\beta)H(2\alpha - 1, 2\beta - 3, 2\delta) - 4\beta(3\alpha - 5\alpha^2 + \beta - 9\alpha\beta + 12\alpha^2\beta)H(2\alpha - 1, 2\beta - 2, 2\delta) \\
& + 4\alpha(8\alpha^2 - 9\alpha + 2)\beta H(2\alpha - 1, 2\beta - 1, 2\delta) + 4\delta(\delta + 1)(\delta + 2\alpha + 6\alpha\delta)H(2\alpha + 1, 2\beta, 2\delta + 3) \\
& - 8\beta\delta(3\delta^2 + 4\delta + 1)H(2\alpha + 1, 2\beta - 1, 2\delta + 3) - 4\alpha\delta(\delta + \alpha(12\delta + 5))H(2\alpha + 1, 2\beta, 2\delta + 2) \\
& - 4\beta\delta(-2 - 5\delta + \beta(12\delta + 5))H(2\alpha + 1, 2\beta - 2, 2\delta + 2) + 4\beta\delta(\delta + 2\alpha(12\delta + 5))H(2\alpha + 1, 2\beta - 1, 2\delta + 2) \\
& + 2\delta^2(3\delta^2 + 8\delta + 5)H(2\alpha + 2, 2\beta, 2\delta + 4) - 8\delta(\delta + 1)(\alpha + \delta + 4\alpha\delta)H(2\alpha + 2, 2\beta, 2\delta + 3) \\
& + 8\beta\delta(4\delta^2 + 5\delta + 1)H(2\alpha + 2, 2\beta - 1, 2\delta + 3) - 4\delta^2(2\delta^2 + 5\delta + 3)H(2\alpha + 3, 2\beta, 2\delta + 4) \\
& + \beta^2(\beta - 1)^2 H^2(\alpha - 1, \beta - 2, \delta) - (2\alpha - 1)\beta^2(\beta - 1)H(\alpha - 1, \beta - 1, \delta)H(\alpha - 1, \beta - 2, \delta) \\
& + \alpha(\alpha - 1)\beta(\beta - 1)H(\alpha - 1, \beta, \delta)H(\alpha - 1, \beta - 2, \delta) - (2\alpha - 1)\beta^2(\beta - 1)H(\alpha - 1, \beta - 2, \delta)H(\alpha - 1, \beta - 1, \delta) \\
& + (\beta - 2\alpha\beta)^2 H^2(\alpha - 1, \beta - 1, \delta) - \alpha(\alpha - 1)(\alpha - 2)\beta H(\alpha - 1, \beta, \delta)H(\alpha - 1, \beta - 1, \delta) \\
& + \alpha(\alpha - 1)\beta(\beta - 1)H(\alpha - 1, \beta - 2, \delta)H(\alpha - 1, \beta, \delta) + \alpha(\alpha - 1)(\alpha - 2)\beta H(\alpha - 1, \beta, \delta)H(\alpha - 1, \beta - 1, \delta) \\
& - \alpha^2(\alpha - 1)^2 H^2(\alpha - 1, \beta, \delta) - 2\beta^2(\beta - 1)\delta H(\alpha, \beta - 1, \delta)H(\alpha - 1, \beta - 2, \delta + 1) \\
& + 2\alpha\beta(\beta - 1)\delta H(\alpha, \beta, \delta + 1)H(\alpha - 1, \beta - 2, \delta) + (4\alpha - 1)\beta^2\delta H(\alpha, \beta - 1, \delta + 1)H(\alpha - 1, \beta - 1, \delta) \\
& - 2\alpha(2\alpha - 1)\beta\delta H(\alpha, \beta, \delta + 1)H(\alpha - 1, \beta - 1, \delta) - 2\alpha(\alpha - 1)\beta\delta H(\alpha, \beta - 1, \delta + 1)H(\alpha - 1, \beta, \delta) \\
& + 2\alpha^2(\alpha - 1)\delta H(\alpha, \beta, \delta + 1)H(\alpha - 1, \beta, \delta) - 4\beta^2(\beta - 1)(\beta - 2)H(\alpha, \beta - 1, \delta)H(\alpha - 1, \beta - 3, \delta) \\
& + 4\alpha\beta(\beta - 1)(\beta - 2)H(\alpha, \beta, \delta)H(\alpha - 1, \beta - 3, \delta) - (\beta - 1)^2\beta^2 H^2(\alpha - 1, \beta - 2, \delta) \\
& + (14\alpha - 9)(\beta - 1)\beta^2 H(\alpha, \beta - 1, \delta)H(\alpha - 1, \beta - 2, \delta) - \alpha(13\alpha - 9)(\beta - 1)\beta H(\alpha, \beta, \delta)H(\alpha - 1, \beta - 2, \delta) \\
& + (2\alpha - 1)(\beta - 1)\beta^2 H(\alpha, \beta - 2, \delta)H(\alpha - 1, \beta - 1, \delta) - (16\alpha^2 - 20\alpha + 5)\beta^2 H(\alpha, \beta - 1, \delta)H(\alpha - 1, \beta - 1, \delta) \\
& + \alpha(14\alpha^2 - 19\alpha + 5)\beta H(\alpha, \beta, \delta)H(\alpha - 1, \beta - 1, \delta) - \alpha(\alpha - 1)\beta(\beta - 1)H(\alpha, \beta - 2, \delta)H(\alpha - 1, \beta, \delta) \\
& + \alpha(6\alpha^2 - 11\alpha + 5)\beta H(\alpha, \beta - 1, \delta)H(\alpha - 1, \beta, \delta) - 5\alpha^2(\alpha - 1)^2 H(\alpha, \beta, \delta)H(\alpha - 1, \beta, \delta) \\
& + \beta(\beta - 1)\beta\delta(\delta + 1)H(\alpha + 1, \beta, \delta + 2)H(\alpha - 1, \beta - 2, \delta) - (2\alpha - 1)\beta\delta(\delta + 1)H(\alpha + 1, \beta, \delta + 2)H(\alpha - 1, \beta - 1, \delta) \\
& + \alpha(\alpha - 1)\delta(\delta + 1)H(\alpha + 1, \beta, \delta + 2)H(\alpha - 1, \beta, \delta) + 4\beta(\beta - 1)(\beta - 2)H(\alpha + 1, \beta, \delta + 1)H(\alpha - 1, \beta - 3, \delta) \\
& + 2\beta^2(\beta - 1)\delta H(\alpha + 1, \beta - 1, \delta + 1)H(\alpha - 1, \beta - 2, \delta) - 2(7\alpha - 4)\beta(\beta - 1)\delta H(\alpha + 1, \beta, \delta + 1)H(\alpha - 1, \beta - 2, \delta) \\
& - 2(2\alpha - 1)\beta^2\delta H(\alpha + 1, \beta - 1, \delta + 1)H(\alpha - 1, \beta - 1, \delta) + 2(8\alpha^2 - 9\alpha + 2)\beta\delta H(\alpha + 1, \beta, \delta + 1)H(\alpha - 1, \beta - 1, \delta) \\
& + 2\alpha(\alpha - 1)\beta\delta H(\alpha + 1, \beta - 1, \delta + 1)H(\alpha - 1, \beta, \delta) - 2\alpha(3\alpha^2 - 5\alpha + 2)\delta H(\alpha + 1, \beta, \delta + 1)H(\alpha - 1, \beta, \delta) \\
& - \beta(\beta - 1)\delta(\delta + 1)H(\alpha + 2, \beta, \delta + 2)H(\alpha - 1, \beta - 2, \delta) + (2\alpha - 1)\beta\delta(\delta + 1)H(\alpha + 2, \beta, \delta + 2)H(\alpha - 1, \beta - 1, \delta) \\
& - \alpha(\alpha - 1)\delta(\delta + 1)H(\alpha + 2, \beta, \delta + 2)H(\alpha - 1, \beta, \delta) - 2\beta^2(\beta - 1)\delta H(\alpha - 1, \beta - 2, \delta)H(\alpha, \beta - 1, \delta + 1) \\
& + 2(2\alpha - 1)\beta^2\delta H(\alpha - 1, \beta - 1, \delta)H(\alpha, \beta - 1, \delta + 1) - \alpha(\alpha - 1)\beta\delta H(\alpha - 1, \beta, \delta)H(\alpha, \beta - 1, \delta + 1) \\
& + 2\alpha\beta(\beta - 1)\delta H(\alpha - 1, \beta - 2, \delta)H(\alpha, \beta, \delta + 1) - 2\alpha(2\alpha - 1)\beta\delta H(\alpha - 1, \beta - 1, \delta)H(\alpha, \beta, \delta + 1) \\
& + 2\alpha^2(\alpha - 1)\delta H(\alpha - 1, \beta, \delta)H(\alpha, \beta, \delta + 1) + 4\beta^2\delta^2 H^2(\alpha, \beta - 1, \delta + 1) - 4\alpha\beta\delta^2 H(\alpha, \beta, \delta + 1)H(\alpha, \beta - 1, \delta + 1) \\
& - 4\alpha\beta\delta^2 H(\alpha, \beta - 1, \delta + 1)H(\alpha, \beta, \delta + 1) + 4\alpha^2\delta^2 H^2(\alpha, \beta, \delta + 1) + 12\beta^2(\beta - 1)\delta H(\alpha, \beta - 1, \delta)H(\alpha, \beta - 2, \delta + 1) \\
& - 12\alpha\beta(\beta - 1)\delta^2 H(\alpha, \beta, \delta)H(\alpha, \beta - 2, \delta + 1) + 2\beta^2(\beta - 1)\delta H(\alpha, \beta - 2, \delta)H(\alpha, \beta - 1, \delta + 1)
\end{aligned}$$

$$\begin{aligned}
& -2(14\alpha - 3)\beta^2\delta H(\alpha, \beta - 1, \delta)H(\alpha, \beta - 1, \delta + 1) + 2\alpha(13\alpha - 3)\beta\delta H(\alpha, \beta, \delta)H(\alpha, \beta - 1, \delta + 1) \\
& - 2\alpha\beta(\beta - 1)\delta H(\alpha, \beta - 2, \delta)H(\alpha, \beta, \delta + 1) + 2\alpha(8\alpha - 3)\beta\delta H(\alpha, \beta - 1, \delta)H(\alpha, \beta, \delta + 1) \\
& - 2\alpha^2(7\alpha - 3)\delta H(\alpha, \beta, \delta)H(\alpha, \beta, \delta + 1) - 2\beta\delta^2(\delta + 1)H(\alpha + 1, \beta, \delta + 2)H(\alpha, \beta - 1, \delta + 1) \\
& + 2\alpha\delta^2(\delta + 1)H(\alpha + 1, \beta, \delta + 2)H(\alpha, \beta, \delta + 1) - 12\beta(\beta - 1)\delta^2(\delta + 1)H(\alpha + 1, \beta, \delta + 1)H(\alpha, \beta - 2, \delta + 1) \\
& - 4\beta^2\delta^2H(\alpha + 1, \beta - 1, \delta + 1)H(\alpha, \beta - 1, \delta + 1) - 4\beta\delta^2H(\alpha + 1, \beta, \delta + 1)H(\alpha, \beta - 1, \delta + 1) \\
& + 28\alpha\beta\delta^2H(\alpha + 1, \beta, \delta + 1)H(\alpha, \beta - 1, \delta + 1) + 4\alpha\beta\delta^2H(\alpha + 1, \beta - 1, \delta + 1)H(\alpha, \beta, \delta + 1) \\
& - 4\alpha(4\alpha - 1)\delta H(\alpha + 1, \beta, \delta + 1)H(\alpha, \beta, \delta + 1) + 2\beta\delta^2(\delta + 1)H(\alpha + 2, \beta, \delta + 2)H(\alpha, \beta - 1, \delta + 1) \\
& - 2\alpha\delta^2(\delta + 1)H(\alpha + 2, \beta, \delta + 2)H(\alpha, \beta, \delta + 1) - (\beta - 1)^2\beta^2H(\alpha - 1, \beta - 2, \delta)H(\alpha, \beta - 2, \delta) \\
& + (2\alpha - 1)(\beta - 1)\beta^2H(\alpha - 1, \beta - 1, \delta)H(\alpha, \beta - 2, \delta) - \alpha(\alpha - 1)\beta(\beta - 1)H(\alpha - 1, \beta, \delta)H(\alpha, \beta - 1, \delta) \\
& + (2\alpha - 1)(\beta - 1)\beta^2H(\alpha - 1, \beta - 2, \delta)H(\alpha, \beta - 1, \delta) - (\beta - 2\alpha\beta)^2H(\alpha - 1, \beta - 1, \delta)H(\alpha, \beta - 1, \delta) \\
& + \alpha(2\alpha^2 - 3\alpha + 1)\beta H(\alpha - 1, \beta, \delta)H(\alpha, \beta - 1, \delta) - \alpha(\alpha - 1)\beta(\beta - 1)H(\alpha - 1, \beta - 2, \delta)H(\alpha, \beta, \delta) \\
& + \alpha(2\alpha^2 - 3\alpha + 1)\beta H(\alpha - 1, \beta - 1, \delta)H(\alpha, \beta, \delta) - \alpha^2(\alpha - 1)^2H(\alpha - 1, \beta, \delta)H(\alpha, \beta, \delta) \\
& + 2\beta^2(\beta - 1)\delta H(\alpha, \beta - 1, \delta + 1)H(\alpha, \beta - 2, \delta) - 2\alpha\beta(\beta - 1)\delta H(\alpha, \beta, \delta + 1)H(\alpha, \beta - 2, \delta) \\
& - 2(2\alpha - 1)\beta^2\delta H(\alpha, \beta - 1, \delta + 1)H(\alpha, \beta - 1, \delta) + 2\alpha(2\alpha - 1)\beta\delta H(\alpha, \beta, \delta + 1)H(\alpha, \beta - 1, \delta) \\
& + 2\alpha(\alpha - 1)\beta\delta H(\alpha, \beta - 1, \delta + 1)H(\alpha, \beta, \delta) - 2\alpha^2(\alpha - 1)\delta H(\alpha, \beta, \delta + 1)H(\alpha, \beta, \delta) \\
& + 4\beta^2(\beta - 1)(\beta - 2)H(\alpha, \beta - 1, \delta)H(\alpha, \beta - 3, \delta) - 4\alpha\beta(\beta - 1)(\beta - 2)H(\alpha, \beta, \delta)H(\alpha, \beta - 3, \delta) \\
& + 3\beta^2(\beta - 1)^2H^2(\alpha, \beta - 2, \delta) - (24\alpha - 10)\beta^2(\beta - 1)H(\alpha, \beta - 1, \delta)H(\alpha, \beta - 2, \delta) \\
& + \alpha(18\alpha - 10)\beta(\beta - 1)H(\alpha, \beta, \delta)H(\alpha, \beta - 2, \delta) - (24\alpha^2 - 20\alpha + 3)\beta^2H^2(\alpha, \beta - 1, \delta) \\
& - \alpha(28\alpha^2 - 30\alpha + 6)\beta H(\alpha, \beta, \delta)H(\alpha, \beta - 1, \delta) + \alpha^2(7\alpha^2 - 10\alpha + 3)H^2(\alpha, \beta, \delta) \\
& - \beta(\beta - 1)\delta(\delta + 1)H(\alpha + 1, \beta, \delta + 2)H(\alpha, \beta, \delta) + (2\alpha - 1)\beta\delta(\delta + 1)H(\alpha + 1, \beta, \delta + 2)H(\alpha, \beta - 1, \delta) \\
& - \alpha(\alpha - 1)\delta(\delta + 1)H(\alpha + 1, \beta, \delta + 2)H(\alpha, \beta, \delta) - 4\beta(\beta - 1)(\beta - 2)\delta H(\alpha + 1, \beta, \delta + 1)H(\alpha, \beta - 3, \delta) \\
& - 6\beta^2(\beta - 1)\delta H(\alpha + 1, \beta - 1, \delta + 1)H(\alpha, \beta - 2, \delta) + 4(9\alpha - 2)\beta(\beta - 1)\delta H(\alpha + 1, \beta, \delta + 1)H(\alpha, \beta - 2, \delta) \\
& + 6(2\alpha - 1)\beta^2\delta H(\alpha + 1, \beta - 1, \delta + 1)H(\alpha, \beta - 1, \delta + 1) - 2\alpha(12\alpha - 7)\beta\delta H(\alpha + 1, \beta, \delta + 1)H(\alpha, \beta - 1, \delta) \\
& - 6\alpha(\alpha - 1)\beta\delta H(\alpha + 1, \beta - 1, \delta + 1)H(\alpha, \beta, \delta) + 10\alpha^2(\alpha - 1)\delta H(\alpha + 1, \beta, \delta + 1)H(\alpha, \beta, \delta) \\
& + 3\beta(\beta - 1)\delta(\delta + 1)H(\alpha + 2, \beta, \delta + 2)H(\alpha, \beta - 2, \delta) - 3(2\alpha - 1)\beta\delta(\delta + 1)H(\alpha + 2, \beta - 1, \delta + 2)H(\alpha, \beta - 1, \delta) \\
& + 3\alpha(\alpha - 1)\delta(\delta + 1)H(\alpha + 2, \beta, \delta + 2)H(\alpha, \beta, \delta) + \beta(\beta - 1)\delta(\delta + 1)H(\alpha - 1, \beta - 2, \delta)H(\alpha + 1, \beta, \delta + 2) \\
& - (2\alpha - 1)\beta\delta(\delta + 1)H(\alpha - 1, \beta - 1, \delta)H(\alpha + 1, \beta, \delta + 2) + \alpha(\alpha - 1)\delta(\delta + 1)H(\alpha - 1, \beta, \delta)H(\alpha + 1, \beta, \delta + 2) \\
& - 2\beta\delta^2(\delta + 1)H(\alpha, \beta - 1, \delta + 1)H(\alpha + 1, \beta, \delta + 2) + 2\alpha\delta^2(\delta + 1)H(\alpha, \beta, \delta + 1)H(\alpha + 1, \beta, \delta + 2) \\
& - 12\beta^2\delta(\delta + 1)H(\alpha, \beta - 1, \delta)H(\alpha + 1, \beta - 1, \delta + 2) + 12\alpha\beta\delta(\delta - 1)H(\alpha, \beta, \delta)H(\alpha + 1, \beta - 1, \delta + 2) \\
& - \beta(\beta - 1)\delta(\delta + 1)H(\alpha, \beta - 2, \delta)H(\alpha + 1, \beta, \delta + 2) + (14\alpha + 3)\beta\delta(\delta + 1)H(\alpha, \beta - 1, \delta)H(\alpha + 1, \beta, \delta + 2) \\
& - \alpha(13\alpha + 3)\delta(\delta + 1)H(\alpha, \beta, \delta)H(\alpha + 1, \beta, \delta + 2) + \delta^2(\delta + 1)^2H^2(\alpha + 1, \beta, \delta + 2) \\
& + 12\beta\delta^2(\delta + 1)H(\alpha + 1, \beta - 1, \delta + 1)H(\alpha + 1, \beta, \delta + 2) - 2(7\alpha + 2)\delta^2(\delta + 1)H(\alpha + 1, \beta, \delta + 1)H(\alpha + 1, \beta, \delta + 2) \\
& - 2\beta^2(\beta - 1)\delta H(\alpha - 1, \beta - 2, \delta)H(\alpha + 1, \beta - 1, \delta + 1) - 2(2\alpha - 1)\beta^2\delta H(\alpha - 1, \beta - 1, \delta)H(\alpha + 1, \beta - 1, \delta + 1) \\
& + 2\alpha(\alpha - 1)\beta\delta H(\alpha - 1, \beta, \delta)H(\alpha + 1, \beta - 1, \delta + 1) - 2\alpha\beta(\beta - 1)\delta H(\alpha - 1, \beta - 2, \delta)H(\alpha + 1, \beta, \delta + 1) \\
& + 2(2\alpha - 1)\beta\delta H(\alpha - 1, \beta - 1, \delta)H(\alpha + 1, \beta, \delta + 1) - 2\alpha^2(\alpha - 1)\delta H(\alpha - 1, \beta, \delta)H(\alpha + 1, \beta, \delta + 1) \\
& - 4\beta^2\delta^2H(\alpha, \beta - 1, \delta + 1)H(\alpha + 1, \beta - 1, \delta + 1) + 8\alpha\beta\delta^2H(\alpha, \beta - 1, \delta + 1)H(\alpha + 1, \beta - 1, \delta + 1) \\
& - 4\alpha^2\beta^2H(\alpha, \beta, \delta + 1)H(\alpha + 1, \beta, \delta + 1) - 12\beta^2(\beta - 1)\delta H(\alpha, \beta - 1, \delta)H(\alpha + 1, \beta - 2, \delta + 1) \\
& + 12\alpha\beta(\beta - 1)\delta H(\alpha, \beta, \delta)H(\alpha + 1, \beta - 2, \delta + 1) - 6\beta^2(\beta - 1)\delta H(\alpha, \beta - 2, \delta)H(\alpha + 1, \beta - 1, \delta + 1) \\
& + 2(18\alpha - 1)\beta^2\delta H(\alpha, \beta - 1, \delta)H(\alpha + 1, \beta - 1, \delta + 1) - 2\alpha(15\alpha - 1)\beta\delta H(\alpha, \beta, \delta)H(\alpha + 1, \beta - 1, \delta + 1) \\
& + 6\alpha\beta(\beta - 1)\delta H(\alpha, \beta - 2, \delta)H(\alpha + 1, \beta, \delta + 1) - 2\alpha(12\alpha - 1)\beta\delta H(\alpha, \beta - 1, \delta)H(\alpha + 1, \beta, \delta + 1) \\
& + 2\alpha^2(9\alpha - 1)\delta H(\alpha, \beta, \delta)H(\alpha + 1, \beta, \delta + 1) + 2\beta\delta^2(\delta + 1)H(\alpha + 1, \beta, \delta + 2)H(\alpha + 1, \beta - 1, \delta + 1)
\end{aligned}$$

$$\begin{aligned}
& -2\alpha\delta^2(\delta+1)H(\alpha+1, \beta, \delta+2)H(\alpha+1, \beta, \delta+1) + 12\beta(\beta-1)\delta^2H(\alpha+1, \beta, \delta+1)H(\alpha+1, \beta-2, \delta+1) \\
& + 12\beta^2\delta^2H^2(\alpha+1, \beta-1, \delta+1) - 4(9\alpha-1)\beta\delta^2H(\alpha+1, \beta, \delta+1)H(\alpha+1, \beta-1, \delta+1) \\
& - 12\alpha\beta\delta^2H(\alpha+1, \beta-1, \delta+1)H(\alpha+1, \beta, \delta+1) + 4\alpha(6\alpha+1)\delta^2H^2(\alpha+1, \beta, \delta+1) \\
& - 6\beta\delta^2(\delta+1)H(\alpha+2, \beta, \delta+2)H(\alpha+1, \beta-1, \delta+1) + 6\alpha\delta^2(\delta+1)H(\alpha+2, \beta, \delta+2)H(\alpha+1, \beta, \delta+1) \\
& + 4\beta\delta(\delta+1)(\delta+2)H(\alpha, \beta-1, \delta)H(\alpha+2, \beta, \delta+3) - 4\alpha\delta(\delta+1)(\delta+2)H(\alpha, \beta, \delta)H(\alpha+2, \beta, \delta+3) \\
& - 4\delta^2(\delta+1)(\delta+2)H(\alpha+1, \beta, \delta+1)H(\alpha+2, \beta, \delta+3) - \beta(\beta-1)\delta(\delta+1)H(\alpha-1, \beta-2, \delta)H(\alpha+2, \beta, \delta+2) \\
& + (2\alpha-1)\beta\delta(\delta+1)H(\alpha-1, \beta-1, \delta)H(\alpha+2, \beta, \delta+2) - \alpha(\alpha-1)\delta(\delta+1)H(\alpha-1, \beta, \delta)H(\alpha+2, \beta, \delta+2) \\
& + 2\beta\delta^2(\delta+1)H(\alpha, \beta-1, \delta+1)H(\alpha+2, \beta, \delta+2) - 2\alpha\delta^2(\delta+1)H(\alpha, \beta, \delta+1)H(\alpha+2, \beta, \delta+2) \\
& + 12\beta^2\delta(\delta+1)H(\alpha, \beta-1, \delta)H(\alpha+2, \beta-1, \delta+2) - 12\alpha\beta\delta(\delta+1)H(\alpha, \beta, \delta)H(\alpha+2, \beta-1, \delta+2) \\
& + 3\beta(\beta-1)\delta(\delta+1)H(\alpha, \beta-2, \delta)H(\alpha+2, \beta, \delta+2) - (18\alpha+5)\beta\delta(\delta+1)H(\alpha, \beta-1, \delta)H(\alpha+2, \beta, \delta+2) \\
& + 5\alpha(3\alpha+1)\delta(\delta+1)H(\alpha, \beta, \delta)H(\alpha+2, \beta, \delta+2) - \delta^2(\delta+1)^2H(\alpha+1, \beta, \delta+2)H(\alpha+2, \beta, \delta+2) \\
& - 12\beta\delta^2(\delta+1)H(\alpha+1, \beta, \delta+1)H(\alpha+2, \beta-1, \delta+2) - 6\beta\delta^2(\delta+1)H(\alpha+1, \beta-1, \delta+1)H(\alpha+2, \beta, \delta+2) \\
& + 2(9\alpha+4)\delta^2(\delta+1)H(\alpha+1, \beta, \delta+1)H(\alpha+2, \beta, \delta+2) + 3\delta^2(\delta+1)^2H^2(\alpha+2, \beta, \delta+2) \\
& - 4\beta\delta(\delta+1)(\delta+2)H(\alpha, \beta-1, \delta)H(\alpha+3, \beta, \delta+3) + 4\alpha\delta(\delta+1)(\delta+2)H(\alpha, \beta, \delta)H(\alpha+3, \beta, \delta+3) \\
& + 4\delta^2(\delta+1)(\delta+2)H(\alpha+1, \beta, \delta+1)H(\alpha+3, \beta, \delta+3)
\end{aligned}$$

$$\begin{aligned}
E(b^3) = & 3\alpha\delta(-\delta+3\alpha(6\delta+1))H(3\alpha, 3\beta, 3\delta+2) + 9\beta\delta(\beta-\delta+6\beta\delta)H(3\alpha, 3\beta-2, 3\delta+2) \\
& - 3\beta\delta(-\delta+6\alpha(1+6\delta))H(3\alpha, 3\beta-1, 3\delta+2) + 30\beta^2(2\beta-1)\delta H(3\alpha, \beta-1, 3\delta+1) \\
& - 12\alpha^2(5\alpha-1)\delta H(3\alpha, 3\beta, 3\delta+1) - 6\beta(-2\beta+5\alpha(6\beta-1))\delta H(3\alpha, 3\beta-2, 3\delta+1) \\
& + 12\alpha(15\alpha-2)\beta\delta H(3\alpha, 3\beta-1, 3\delta+1) - \beta^2(-7\beta+18\alpha(2\beta-1))H(3\alpha-2, \beta-1, 3\delta) \\
& + \alpha^3(9\alpha-7)H(3\alpha-2, 3\beta, 3\delta) + \beta^3(9\beta-9)H(3\alpha-2, 3\beta-4, 3\delta) \\
& + 3\alpha\beta(-7\beta+3\alpha(6\beta-1))H(3\alpha-2, 3\beta-2, 3\delta) - 3\alpha^2(12\alpha-7)\beta H(3\alpha-2, 3\beta-1, 3\delta) \\
& - 18\beta^2(2\beta-1)\delta H(3\alpha-1, 3\beta-3, 3\delta+1) + 12\alpha^2(3\alpha-1)\delta H(3\alpha-1, 3\beta, 3\delta+1) \\
& + 6\beta(-2\beta+3\alpha(6\beta-1))\delta H(3\alpha-1, 3\beta-2, 3\delta+1) - 12\alpha(9\alpha-2)\beta\delta H(3\alpha-1, 3\beta-1, 3\delta+1) \\
& + 3\beta^2(-3\beta+10\alpha(-1+2\beta))H(3\alpha-1, 3\beta-3, 3\delta) - 3\alpha^3(-3+5\alpha)H(3\alpha-1, 3\beta, 3\delta) \\
& - \beta^3(15\beta-15)H(3\alpha-1, 3\beta-4, 3\delta) - 3\alpha\beta(-9\beta+5\alpha(-1+6\beta))H(3\alpha-1, 3\beta-2, 3\delta) \\
& + 3\alpha^2(-9+20\alpha)\beta H(3\alpha-1, 3\beta-1, 3\delta) - 3\alpha\delta(\delta+5\alpha(1+6\delta))H(3\alpha+1, 3\beta, 3\delta+2) \\
& - 15\beta\delta(\beta-\delta+6\beta\delta)H(3\alpha+1, 3\beta-2, 3\delta+2) + 3\beta\delta(\delta+10\alpha(1+6\delta))H(3\alpha+1, 3\beta-1, 3\delta+2) \\
& + 2\delta^2(\delta+9\alpha(2\delta+1))H(3\alpha+1, 3\beta, 3\delta+3) - 18\beta\delta^2(2\delta+1)H(3\alpha+1, 3\beta-1, 3\delta+3) \\
& + 9\delta^3(\delta+1)H(3\alpha+2, 3\beta, 3\delta+4) - 6\delta^2(\delta+5\alpha(2\delta+1))H(3\alpha+2, 3\beta, 3\delta+3) \\
& + 30\beta\delta^2(2\delta+1)H(3\alpha+2, 3\beta-1, 3\delta+3) - 15\delta^3(1+\delta)H(3\alpha+3, 3\beta, 3\delta+4) \\
& - 4\beta^3H(\alpha+2, \beta-1, \delta)H^2(\alpha-1, \beta-1, \delta) + 4\alpha\beta^2H(\alpha+2, \beta, \delta)H^2(\alpha-1, \beta-1, \delta) \\
& + 4\alpha\beta^2H(\alpha+2, \beta-1, \delta)H(\alpha-1, \beta, \delta)H(\alpha-1, \beta-1, \delta) - 4\alpha^2\beta H(\alpha+2, \beta, \delta)H(\alpha-1, \beta, \delta)H(\alpha-1, \beta-1, \delta) \\
& + 4\alpha\beta^2H(\alpha+2, \beta-1, \delta)H(\alpha-1, \beta-1, \delta)H(\alpha-1, \beta, \delta) - 4\alpha^2\beta H(\alpha+2, \beta, \delta)H(\alpha-1, \beta-1, \delta)H(\alpha-1, \beta, \delta) \\
& - 4\alpha^2\beta H(\alpha+2, \beta-1, \delta)H^2(\alpha-1, \beta, \delta) + 4\alpha^3H(\alpha+2, \beta, \delta)H^2(\alpha-1, \beta, \delta) \\
& + 4\beta^2\delta H(\alpha+3, \beta, \delta+1)H^2(\alpha-1, \beta-1, \delta) - 4\alpha\beta\delta H(\alpha+3, \beta, \delta+1)H(\alpha-1, \beta, \delta)H(\alpha-1, \beta-1, \delta) \\
& - 4\alpha\beta\delta H(\alpha+3, \beta, \delta+1)H(\alpha-1, \beta-1, \delta)H(\alpha-1, \beta, \delta) + 4\alpha^2\delta H(\alpha+3, \beta, \delta+1)H(\alpha-1, \beta, \delta)H(\alpha-1, \beta, \delta)
\end{aligned}$$

$$\begin{aligned}
& -9\alpha(-1+2\alpha)\beta\delta H(\alpha, \beta, \delta)H(\alpha+1, \beta, \delta+1)H(\alpha, \beta-1, \delta) - 9(-1+\alpha)\alpha\beta\delta H(\alpha, \beta-1, \delta)H(\alpha+1, \beta, \delta+1)H(\alpha, \beta, \delta) \\
& + 9(-1+\alpha)\alpha^2\delta H^2(\alpha, \beta, \delta)H(\alpha+1, \beta, \delta+1) + 9(-1+\beta)\beta\delta^2 H^2(\alpha+1, \beta, \delta+1)H(\alpha, \beta-2, \delta) \\
& - 9(-1+2\alpha)\beta\delta^2 H^2(\alpha+1, \beta, \delta+1)H(\alpha, \beta-1, \delta) + 9(-1+\alpha)\alpha\delta^2 H^2(\alpha+1, \beta, \delta+1)H(\alpha, \beta, \delta) \\
& - 3\beta^2\delta(1+\delta)H^2(\alpha, \beta-1, \delta)H(\alpha+1, \beta, \delta+2) + 3\alpha\beta\delta(1+\delta)H(\alpha, \beta, \delta)H(\alpha, \beta-1, \delta)H(\alpha+1, \beta, \delta+2) \\
& + 3\alpha\beta\delta(1+\delta)H(\alpha, \beta-1, \delta)H(\alpha, \beta, \delta)H(\alpha+1, \beta, \delta+2) - 3\alpha^2\delta(1+\delta)H^2(\alpha, \beta, \delta)H(\alpha+1, \beta, \delta+2) \\
& + 3\beta\delta^2(1+\delta)H(\alpha+1, \beta, \delta+1)H(\alpha, \beta-1, \delta)H(\alpha+1, \beta, \delta+2) - 3\alpha\delta^2(1+\delta)H(\alpha+1, \beta, \delta+1)H(\alpha, \beta, \delta)H(\alpha+1, \beta, \delta+2) \\
& + 3\beta\delta^2(1+\delta)H(\alpha, \beta-1, \delta)H(\alpha+1, \beta, \delta+1)H(\alpha+1, \beta, \delta+2) - 3\alpha\delta^2(1+\delta)H(\alpha, \beta, \delta)H(\alpha+1, \beta, \delta+1)H(\alpha+1, \beta, \delta+2) \\
& - 3\delta^3(1+\delta)H^2(\alpha+1, \beta, \delta+1)H(\alpha+1, \beta, \delta+2) - 18\beta^3\delta H^2(\alpha, \beta-1, \delta)H(\alpha+1, \beta-1, \delta+1) \\
& + 18\alpha\beta^2\delta H(\alpha, \beta, \delta)H(\alpha, \beta-1, \delta)H(\alpha+1, \beta-1, \delta+1) + 18\alpha\beta^2\delta H(\alpha, \beta-1, \delta)H(\alpha, \beta, \delta)H(\alpha+1, \beta-1, \delta+1) \\
& - 18\alpha^2\beta\delta H^2(\alpha, \beta, \delta)H(\alpha+1, \beta-1, \delta+1) + 18\alpha\beta^2\delta H^2(\alpha, \beta-1, \delta)H(\alpha+1, \beta, \delta+1) \\
& - 18\alpha^2\beta\delta H(\alpha, \beta, \delta)H(\alpha, \beta-1, \delta)H(\alpha+1, \beta, \delta+1) - 18\alpha^2\beta\delta H(\alpha, \beta-1, \delta)H(\alpha, \beta, \delta)H(\alpha+1, \beta, \delta+1) \\
& + 18\alpha^3\delta H^2(\alpha, \beta, \delta)H(\alpha+1, \beta, \delta+1) + 18\beta^2\delta^2 H(\alpha+1, \beta, \delta+1)H(\alpha, \beta-1, \delta)H(\alpha+1, \beta-1, \delta+1) \\
& - 18\alpha\beta\delta^2 H(\alpha+1, \beta, \delta+1)H(\alpha, \beta, \delta)H(\alpha+1, \beta-1, \delta+1) - 18\alpha\beta\delta^2 H(\alpha+1, \beta, \delta+1)H(\alpha, \beta-1, \delta)H(\alpha+1, \beta, \delta+1) \\
& + 18\alpha^2\delta^2 H(\alpha+1, \beta, \delta+1)H(\alpha, \beta, \delta)H(\alpha+1, \beta, \delta+1) + 18\beta^2\delta^2 H(\alpha, \beta-1, \delta)H(\alpha+1, \beta, \delta+1)H(\alpha+1, \beta-1, \delta+1) \\
& - 18\alpha\beta\delta^2 H(\alpha, \beta, \delta)H(\alpha+1, \beta, \delta+1)H(\alpha+1, \beta-1, \delta+1) - 18\alpha\beta\delta^2 H(\alpha, \beta-1, \delta)H^2(\alpha+1, \beta, \delta+1) \\
& + 18\alpha^2\delta^2 H(\alpha, \beta, \delta)H(\alpha+1, \beta, \delta+1)H(\alpha+1, \beta, \delta+1) - 18\beta\delta^3 H^2(\alpha+1, \beta, \delta+1)H(\alpha+1, \beta-1, \delta+1) \\
& + 18\alpha\delta^3 H^3(\alpha+1, \beta, \delta+1) + 9\beta^2\delta(1+\delta)H^2(\alpha, \beta-1, \delta)H(\alpha+2, \beta, \delta+2) \\
& - 9\alpha\beta\delta(1+\delta)H(\alpha, \beta, \delta+2)H(\alpha, \beta-1, \delta)H(\alpha+2, \beta, \delta+2) - 9\alpha\beta\delta(1+\delta)H(\alpha, \beta-1, \delta)H(\alpha, \beta, \delta)H(\alpha+2, \beta, \delta+2) \\
& + 9\alpha^2\delta(1+\delta)H^2(\alpha, \beta, \delta)H(\alpha+2, \beta, \delta+2) - 9\beta\delta^2(1+\delta)H(\alpha+1, \beta, \delta+1)H(\alpha, \beta-1, \delta)H(\alpha+2, \beta, \delta+2) \\
& + 9\alpha\delta^2(1+\delta)H(\alpha+1, \beta, \delta+1)H(\alpha, \beta, \delta)H(\alpha+2, \beta, \delta+1) - 9\beta\delta^2(1+\delta)H(\alpha, \beta-1, \delta)H(\alpha+1, \beta, \delta+1) \\
& + 9\alpha\delta^2(1+\delta)H(\alpha, \beta, \delta)H(\alpha+1, \beta, \delta+1)H(\alpha+2, \beta, \delta+2) + 9\delta^3(1+\delta)H(\alpha+1, \beta, \delta+1)H(\alpha+1, \beta, \delta+1)
\end{aligned}$$

A partir das expressões de $E(b)$, $E(b^2)$, $E(b^3)$ e $E(b^4)$ é possível obter o valor esperado a variância, assimetria e curtose de $h(\alpha, \beta, \delta)$, respectivamente, como segue:

$$E[h(\alpha, \beta, \delta)] = E(b) + H(\alpha, \beta, \delta) \quad (4.8)$$

$$\begin{aligned} \text{Var}[h(\alpha, \beta, \delta)] &= E\{[h(\alpha, \beta, \delta) - E(h(\alpha, \beta, \delta))]^2\} \\ \text{Var}[h(\alpha, \beta, \delta)] &= E(b^2) - E^2(b) \end{aligned} \quad (4.9)$$

$$\begin{aligned} A[h(\alpha, \beta, \delta)] &= E\{[h(\alpha, \beta, \delta) - E(h(\alpha, \beta, \delta))]^3\} \\ A[h(\alpha, \beta, \delta)] &= \frac{E(b^3) - 3E(b)E(b^2) + 2E^3(b)}{\text{Var}[h(\alpha, \beta, \delta)]^{3/2}} \end{aligned} \quad (4.10)$$

$$\begin{aligned} K[h(\alpha, \beta, \delta)] &= E\{[h(\alpha, \beta, \delta) - E(h(\alpha, \beta, \delta))]^4\} \\ K[h(\alpha, \beta, \delta)] &= \frac{E(b^4) - 4E(b)E(b^3) + 6E^2(b) - 3E^4(b)}{\text{Var}[h(\alpha, \beta, \delta)]^2} \end{aligned} \quad (4.11)$$

em que as medidas de assimetria e curtose são controladas pelos parâmetros α, β, δ .

4.2 Assimetria e curtose de $h(\alpha, \beta, \delta)$

As tabelas 1 e 2 apresentam os momentos para $h(\alpha, \beta, \delta)$ para uma série de diferentes valores de α , β e δ variando na região de aceitação como dado em Baczkowski *et al.* (1997, 1998), obtidos ao utilizar as fórmulas apresentadas na seção anterior, no caso em que $s = 10$ e $n = 1000$ para o modelo de broken-stick de MacArthur (1957), adotado como um modelo representativo com diferentes abundâncias relativas e o modelo equiprovável, respectivamente.

De uma forma geral, ao aplicar o modelo de broken-stick para quantificar a abundância entre espécies, nota-se que a medida que os parâmetros β e δ crescem o índice de diversidade, $H(\alpha, \beta, \delta)$, também cresce, ou seja, tais parâmetros apresentam uma relação diretamente proporcional com o índice. O mesmo não acontece ao observar o parâmetro α que apresenta uma relação inversamente proporcional ao índice já que o mesmo decresce quando α cresce.

Ainda nessa tabela, é possível caracterizar a distribuição do índice de diversidade a partir dos momentos de $h(\alpha, \beta, \delta)$ em que, a aproximação normal não é vista como a mais aconselhável já que além de apresentar um comportamento assimétrico, a curtose, para os valores atribuídos a α, β, δ , evidenciam que a distribuição dos dados possui caudas pesadas.

Tabela 1. Valores de $H(\alpha, \beta, \delta)$ e momentos de $h(\alpha, \beta, \delta)$ com $s = 10$, $n = 1000$ para o modelo broken-stick.

α	β	δ	$H(\alpha, \beta, \delta)$	$E(h)$	$Var(h) \times 10^6$	$A(h) \times 10^8$	$K(h) \times 10^{10}$
0,5	0,0	1,0	3,3808	3,3775	32,3630	-0,0008	0,1160
0,5	0,5	2,0	5,7280	5,7219	165,1248	-0,0007	0,1107
1,0	-2,0	2,0	0,6099	0,6180	258,2401	-0,0008	0,6181
1,0	-1,5	0,5	0,5039	0,5070	22,7916	-0,0009	0,5130
1,0	-1,0	0,5	0,6383	0,6410	13,0031	-0,0009	0,4407
1,0	-0,5	1,0	0,9244	0,9272	15,6271	-0,0009	0,3634
1,0	0,5	0,5	1,5114	1,5102	3,3522	-0,0009	0,0262
1,0	1,0	0,0	1,9723	1,9678	18,5119	-0,0004	0,0769
1,0	1,5	2,0	3,9055	3,8987	83,8917	-0,0006	0,0681
1,5	-2,0	1,0	0,2192	0,2224	31,7728	-0,0010	0,7352
1,5	-1,5	1,0	0,2626	0,2656	23,1275	-0,0010	0,6894
1,5	-1,0	2,0	0,4183	0,4230	59,1860	-0,0010	0,7138
1,5	-0,5	1,5	0,4483	0,4515	4429,4797	-0,0012	0,0537
1,5	0,0	1,0	0,4975	0,4995	4,9593	-0,0010	0,3642
2,0	-2,0	0,5	0,0926	0,0942	7,4460	-0,0012	0,8739
2,0	-1,5	1,5	0,1465	0,1490	19,9053	-0,0011	0,8380
2,0	-1,0	1,5	0,1732	0,1755	14,8536	-0,0011	0,8042
2,0	-0,5	0,5	0,1575	0,1587	2,4741	-0,0012	0,6173
2,0	0,0	0,0	0,1701	0,1710	0,5170	-0,0011	0,3519

Assim como descrito na tabela 1, nota-se que os parâmetros β e δ apresentam um relação diretamente proporcional, enquanto o parâmetro α apresenta uma relação inversamente proporcional ao índice quando se aplica o modelo equiprovável para obter a abundância entre espécies.

Nota-se que independente de obter a abundância entre espécies através doo modelo de broken-stick ou do modelo equiprovável a aproximação normal não é vista como a mais aconselhável para caracterizar a distribuição do índice de diversidade a partir dos momentos de $h(\alpha, \beta, \delta)$, já que além de apresentar um comportamento assimétrico os valores da curtose evidenciam que a distribuição dos dados possui caudas pesadas, para os valores atribuídos a α, β, δ .

Tabela 2. Valores de $H(\alpha, \beta, \delta)$ e momentos de $h(\alpha, \beta, \delta)$ com $s = 10$, $n = 1000$ para o modelo equiprovável.

α	β	δ	$H(\alpha, \beta, \delta)$	$E(h)$	$Var(h) \times 10^6$	$A(h) \times 10^8$	$K(h) \times 10^{10}$
0,5	0,0	1,0	3,5140	0,3512	1,5039	-20,7133	0,1329
0,5	0,5	2,0	5,9240	0,5921	6,2335	-14,3620	0,0656
1,0	-2,0	2,0	0,2330	0,0236	0,7461	-24,9508	0,8375
1,0	-1,5	0,5	0,3020	0,0304	0,3376	-61,3611	0,8921
1,0	-1,0	0,5	0,4580	0,0460	0,3310	-110,5478	0,8301
1,0	-0,5	1,0	0,7320	0,0734	0,4771	-148,4061	0,6404
1,0	0,5	0,5	0,1600	0,1598	0,3401	-12,9635	0,2693
1,0	1,0	0,0	2,3026	0,2298	1,6735	33,2862	0,8138
1,0	1,5	2,0	4,3136	0,3486	6,6339	9,6725	0,2811
1,5	-2,0	1,0	0,0663	0,0068	0,1035	-22,3470	1,0302
1,5	-1,5	1,0	0,1006	0,0102	0,1460	-29,3620	1,0298
1,5	-1,0	2,0	0,1696	0,0172	0,3485	-31,2036	0,9836
1,5	-0,5	1,5	0,2441	0,0246	3,8597	-16,5380	0,9412
1,5	0,0	1,0	0,3514	0,0353	0,2197	-120,0551	0,9532
2,0	-2,0	0,5	0,0199	0,0020	0,0184	-17,1348	1,1119
2,0	-1,5	1,5	0,0335	0,0034	0,0454	-17,9861	1,0873
2,0	-1,0	1,5	0,0509	0,0052	0,0686	-22,2146	1,0914
2,0	-0,5	0,5	0,0695	0,0070	0,0550	-38,2523	1,1536
2,0	0,0	0,0	0,1000	0,0101	0,0450	-77,9739	1,2744

5 APLICAÇÕES

Para os dados coletados no Parque Estadual Mata dos Godoy, comprovou-se a existência de uma relação diretamente proporcional entre os parâmetros β e δ e o índice de diversidade. Nota-se ainda que a aproximação normal não é vista como a distribuição mais adequada aos dados devido as caudas pesadas, caracterizada pela curtose.

Tabela 3. Valores de $H(\alpha, \beta, \delta)$ e momentos de $h(\alpha, \beta, \delta)$ com $s = 65$, $n = 970$ para o modelo equiprovável.

α	β	δ	$H(\alpha, \beta, \delta)$	$E(h)$	$Var(h) \times 10^6$	$A(h) \times 10^8$	$K(h) \times 10^{10}$
0,5	0,0	1,0	0,1260	-0,001001	6,6493	-0,1137	0,0097
0,5	0,5	2,0	0,2615	-0,002134	37,6217	-0,0417	0,0075
1,0	-2,0	2,0	0,0009	0,000028	0,0053	-1,0249	0,0098
1,0	-1,5	0,5	0,0018	0,000036	0,0083	-1,3213	0,0107
1,0	-1,0	0,5	0,0037	0,000046	0,0134	-1,6676	0,0109
1,0	-0,5	1,0	0,0077	0,000050	0,0215	-2,1404	0,0078
1,0	0,5	0,5	0,0317	-0,000127	0,2770	-0,6869	0,0036
1,0	1,0	0,0	0,0643	-0,000513	2,7841	-0,1385	0,0061
1,0	1,5	2,0	0,1354	-0,001248	21,1986	-0,0383	0,0046
1,5	-2,0	1,0	0,0001	0,000008	0,0004	-1,7859	0,0112
1,5	-1,5	1,0	0,0002	0,000013	0,0011	-1,2928	0,0113
1,5	-1,0	2,0	0,0005	0,000023	0,0032	-0,8773	0,0110
1,5	-0,5	1,5	0,0013	0,000046	0,0144	-0,5292	0,0099
1,5	0,0	1,0	0,0019	0,000051	0,0163	-0,7144	0,0111
2,0	-2,0	0,5	0,0000	0,000002	0,0000	-4,5397	0,0115
2,0	-1,5	1,5	0,0000	0,000003	0,0001	-2,7438	0,0115
2,0	-1,0	1,5	0,0001	0,000006	0,0002	-1,8064	0,0115
2,0	-0,5	0,5	0,0001	0,000009	0,0005	-1,3253	0,0117
2,0	0,0	0,0	0,0002	0,000016	0,0015	-0,9808	0,0118

Inicialmente comprova-se a relação diretamente proporcional entre os parâmetros β e δ e o índice de diversidade, assim como ocorrido nas tabelas 1, 2 e 3. Ainda nessa tabela, ao caracterizar a distribuição do índice de diversidade com base nos dados que compõe um estudo desenvolvido no parque fenológico do Campo Experimental da Embrapa Amazônia Ocidental nota-se que a aproximação normal não é vista como a distribuição mais adequada aos dados devido as caudas pesadas e comportamento assimétrico.

Tabela 4. Valores de $H(\alpha, \beta, \delta)$ e momentos de $h(\alpha, \beta, \delta)$ com $s = 99$, $n = 240$ para o modelo equiprovável.

α	β	δ	$H(\alpha, \beta, \delta)$	$E(h)$	$Var(h) \times 10^6$	$A(h) \times 10^8$	$K(h) \times 10^{10}$
0,5	0,0	1,0	0,1260	-0,005892	6,6493	-0,1137	0,0097
0,5	0,5	2,0	0,2615	-0,013358	37,6217	-0,0417	0,0075
1,0	-2,0	2,0	0,0009	0,000079	0,0053	-1,0249	0,0098
1,0	-1,5	0,5	0,0018	0,000111	0,0083	-1,3213	0,0107
1,0	-1,0	0,5	0,0037	0,000150	0,0134	-1,6676	0,0109
1,0	-0,5	1,0	0,0077	0,000167	0,0215	-2,1404	0,0078
1,0	0,5	0,5	0,0317	-0,000523	0,2770	-0,6869	0,0036
1,0	1,0	0,0	0,0643	-0,002206	2,7841	-0,1385	0,0061
1,0	1,5	2,0	0,1354	-0,005987	21,1986	-0,0383	0,0046
1,5	-2,0	1,0	0,0001	0,000020	0,0004	-1,7859	0,0112
1,5	-1,5	1,0	0,0002	0,000035	0,0011	-1,2928	0,0113
1,5	-1,0	2,0	0,0005	0,000063	0,0032	-0,8773	0,0110
1,5	-0,5	1,5	0,0010	0,000127	0,0144	-0,4286	0,0099
1,5	0,0	1,0	0,0019	0,000165	0,0163	-0,7144	0,0111
2,0	-2,0	0,5	0,0000	0,000004	0,0000	-4,5397	0,0115
2,0	-1,5	1,5	0,0000	0,000007	0,0001	-2,7438	0,0115
2,0	-1,0	1,5	0,0001	0,000013	0,0002	-1,8064	0,0115
2,0	-0,5	0,5	0,0001	0,000024	0,0005	-1,3253	0,0117
2,0	0,0	0,0	0,0002	0,000042	0,0015	-0,9808	0,0118

6 CONSIDERAÇÕES FINAIS

Os momentos de $h(\alpha, \beta, \delta)$ apresentados nessa dissertação estende os resultados apresentados em Baczkowski *et al.* (2000) e Bowman *et al.* (1971) para uma classe de índices de diversidade mais geral. Tais resultados são tão consistentes quanto os trabalhos que lidam com populações reais como em Heip & Engels (1974), que examina o habitat aquático, principalmente quando aborda-se o índice de Shannon.

Ao caracterizar a distribuição do índice proposto através de momentos, nota-se que a aproximação normal não é a mais adequada, já que ao quantificar a abundância de espécies, tanto pelo modelo de broken-stick quanto pelo modelo equiprovável, é possível observar a ocorrência de assimetria e de caudas pesadas, isto ao estudar encontros intra-específicos para a amostra de valores apresentados nas tabelas 1 e 2, respectivamente.

O mesmo acontece ao tratar dados reais, sejam os dados referentes a estrutura de espécies arbóreas em uma área alagável do Município de Londrina, sejam os dados referentes a associação florística e fitossociologia de espécies arbóreas do Parque Fenológico da Embrapa Amazônia Ocidental, como mostram as tabelas 3 e 4.

Em Baczkowski *et al.* (2000) a distribuição de $h(1, 1, 0)$ pode ser aproximada por uma distribuição Gaussiana, caso haja diferença entre abundância de espécies, no entanto, ao atribuir valores mais gerais para os parâmetros (α, β, δ) , sugere-se a distribuição beta tipo I como sendo a mais apropriada para caracterizar a distribuição do índice proposto. Nesta dissertação, também é possível usar a aproximação Gaussiana como um modelo para a distribuição do índice, no entanto não é aconselhável que se faça isto pelos motivos já citados.

Muito ainda pode ser feito ao aplicar ferramentas estatística em ecologia, a princípio sugere-se estudar os limites de δ tomando como base Baczkowski *et al.* (1998). Recomenda-se ainda um estudo mais aprofundado sobre as inter-relações entre espécies, enfatizando o cenário interespecífico, já que este corresponde a prováveis eventos raros. Ainda, para trabalhos futuros, sugere-se a criação de um índice de diversidade que abranja não somente o número de espécies mas também o número de indivíduos por espécies, atribuindo, com isso, pesos a cada espécie objetivando minimizar o erro de estimação já que ao ponderar a dominância entre indivíduos a estimativa de sua quantificação em uma dada comunidade convergirá para o real valor da diversidade.

REFERÊNCIAS BIBLIOGRÁFICAS

BACZKOWSKI, A.J.; JOANES, D.N.; SHAMIA, G. The distribution of a generalized diversity index due to Good. **Environmental and Ecological Statistics 7**: 329-342, 2000

BACZKOWSKI, A.J. Moments of a generalized diversity index due to Good. Unpublished Internal Report STAT 96/18, December 1996. Department of Statistics, University of Leeds, 1996.

BACZKOWSKI, A.J. (1998) Applications of Good's generalized diversity index. Unpublished Internal Report STAT 98/11, September 1998, Department of Statistics, University of Leeds, 1998.

BACZKOWSKI, A.J.; JOANES, D.N.; SHAMIA, G. Properties of a generalized diversity index. **Journal of Theoretical Biology 188**: 207-213, 1997.

BACZKOWSKI, A.J.; JOANES, D.N.; SHAMIA, G.M. Range of validity of a and b for a generalized diversity index $H(\alpha, \beta)$ due to Good. **Mathematical Biosciences 148**: 115-128, 1998.

BASHARIN, G.P. On a statistical estimate for the entropy of a sequence of independent random variables. In Theory of Probability and its Applications, v.IV, (Translation of Teoriya Veroyatnostei ee Pvmneniya), N. Artin (ed), **Society for Industrial and Applied Mathematics**: 333-336, 1959.

BERGER, W.H.; PARKER, F.L. Diversity of planktonic forminifera in deep-sea sediments. **Science 168**: 1345-1347, 1970.

BHARGAVA, T.N.; UPPULURI, V.R.R. Population disagreement and Gini's index of diversity. **General Systems 21**: 101-103, 1976.

BIANCHINI, E.; POPOLO, R.S.; DIAS, M.C; PIMENTA, J.A. Diversidade e estrutura de espécies arbóreas em área alagável do município de Londrina, Sul do Brasil. **Acta. bot. bras. 17 (3)**: 405-419, 2003.

BOWMAN, K.O. *et al.* Comments on the distribution of indices of diversity. **International Symposium of Statistical Ecology 3**: 315-359, 1971.

CORDEIRO, G.M. **Introdução a teoria assintótica**. Rio de Janeiro: 22º Colóquio Brasileiro de Matemática, 1999.

FISHER, R.A. *et al.* The relation between the number of species and the number of individuals in a sample de an animal population. **Journal Anim Ecol.12**: 42-58, 1943.

GOOD, I.J. The population frequencies of species and the estimation of population parameters. **Biometrika 40**: 237-64, 1953.

GOOD, I.J. Comment on the paper by G.P. Patil and C. Taillie. **Journal of the American Statistical Association 77**: 561-563, 1982.

GREENBERG, S.H. The measurement of linguistic diversity. **Language 32**: 109-115, 1956.

HAIRSTON, N.G. Species abundance and community organization. **Ecology 40**: 404-416, 1959.

HEIP, C.; ENGELS, P. Comparing species diversity and evenness indices. **Journal of the Marine Biological Association of the UK 54**: 559-563, 1974.

HERDAN, G. The mathematical relation between Greenberg's index of linguistic diversity and Yule's characteristic. **Biometrika 45**: 268-270, 1958.

HILL, M.O. Diversity and evenness: a unifying notation and its consequences. **Ecology 54**: 427-432, 1973.

HULBERT, S. H. The nonconcept of species diversity: a critique and alternative parameters. **Ecology 52**: 577-586, 1971.

HUTCHESON, K. A test for comparing diversities based on the Shannon formula. **Journal of Theoretical Biology 29**: 151-154, 1970.

- JOANES, D.N.; GILL, C.A. Interval estimation of a measure of diversity. **Journal of Applied Statistics** **19**: 323-327, 1992.
- JOANES, D.N.; GILL, C.A. Comparing measures of sample skewness and kurtosis. **The Statistician** **47**: 183-189, 1998.
- JOHNSON, N. L.; KOTZ, S. **Encyclopedia of statistical science**, v.2. New York: John Wiley, 1988.
- KHINCHIN, A. I. **Mathematical Foundations of Information Theory**. Dover, New York, 1957.
- KOLASA, J.; STRAYER, D. Patterns of the abundance of species: a comparison of two hierarchical models. **Oikos** **53**: 235-241, 1988.
- MACARTHUR, R.H. On the relative abundance of bird species. **Proceedings of the National Academy of Sciences, USA** **43**: 293-295, 1957.
- MAGURRAN, A.E. **Ecological diversity and its measurements**. Princeton: University Press, 1988.
- MARGALEF, R. La teoría de la información en ecología. **Memórias de la Real Academia de Ciencias y Artes de Barcelona** **32**: 373-449, 1957.
- MARGALEF, R. Information theory in ecology. **General Systems** **3**: 36-71, 1958.
- MARTINS, F. R.; SANTOS, F. A. M. Técnicas usuais de estimativas da biodiversidade. **Holos** **1**: 236-267, 1999.
- MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction to the theory of statistics**. McGraw-Hill, 1974.
- PATIL, G.P.; TAILLIE, C. Diversity as a concept and its measurement. **Journal of the American Statistical Association** **77 (379)**: 548-561, 1982.

PEARSON, E.S.; HARTLEY, H.O. **Biometrika Tables for Statisticians**, Volume I, Cambridge: University Press, 1954.

PEARSON, E.S.; MERRINGTON, M. Tables of the 5% and 0.5% point of Pearson curves (with argument b_1 and b_2) expressed in standard measure. **Biometrika** **38**: 4-10, 1951.

PIELOU, E.C. **Ecological Diversity**. New York: John Wiley & Sons, 1975.

PIELOU, E.C. **Mathematical Ecology**. New York: John Wiley & Sons, 1977.

PINTO-COELHO, R. M. **Fundamentos em Ecologia**. Porto Alegre: Artmed, 2006.

PRESTON, F.W. The commonness and rarity of species. **Ecology** **29**: 254-283, 1948.

RENYI, A. On measures of entropy and information: In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* (Neyman, J., ed) Volume 1, 436-450. Berkeley: University of California, 1961.

SIMPSON, E.H. Measurement of diversity. **Nature** **163**: 688, 1949.

SHANNON, C.E. A mathematical theory of communication. **The Bell System Technical Journal** **27**: 379-423, 1948.

SHANNON, C.E.; WEAVER, W. **The mathematical theory of communication**. Urbana: University of Illinois Press, 1949.

SILVA, K. E. da; MATOS, F. D. de; FERREIRA, M. M. Composição florística e fitossociologia de espécies arbóreas do Parque Fenológico da Embrapa Amazônia Ocidental. **Acta amazonica** **38**: 213-222, 2008.

SUGIHARA, G. Minimal community structure: an explanation of species abundance patterns. **The American Naturalist** **116**: 770-787, 1980.

WALLACE, A.F. **Tropical Nature and outhier Essays**. London: MacMillan, 1875.

WEBB, D.J. The statistics of relative abundance and diversity. **Journal of Theoretical Biology** **43**: 277-292, 1974.

WIKIPEDIA. Diversidade. Informações disponíveis no site <<http://pt.wikipedia.org/wiki/Diversidade>>. Acesso em: 30 set. 2008.

WILLIAMS, C.B. **Patterns in the balance of nature**. London: Academic Press, 1964.

WHITTAKER, R.H. Dominance and diversity in land plant communities. **Science** **147**: 250-260, 1967.

WHITTAKER, R.H. Evolution and measurement of species diversity. **Taxon** **21**: 17-34, 1972.