STATISTICAL ANALYSIS AND DIVERSITY WITH SPECIAL REFERENCE TO BRAZILIAN FISH

CARLOS ARTUR SOBREIRA ROCHA

STATISTICAL ANALYSIS AND DIVERSITY

WITH SPECIAL REFERENCE TO

BRAZILIAN FISH

By

(C) CARLOS ARTUR SOBREIRA ROCHA

A Thesis submitted in partial fulfillment of the requirements for the

degree of Master of Science

Department of Mathematics and Statistics

Memorial University of Newfoundland

December, 1980

St. John's Newfoundland

ABSTRACT

In this thesis, a brief review of various techniques of estimating diversity is given. The use of these techniques is illustrated to the fish data. The data is collected from Brazilian coast utilizing sampling techniques. In order to reduce the bias, the jackknife method is used for estimating the diversity. Furthermore, it has been shown that lognormal distribution is a suitable model to describe species abundance in certain cases.

to:

Iracema

Jeritza

Carlos Arthur Filho

ACKNOWLEDGEMENTS

The author wishes to express his gratitude to Professor Rajendra Jain, his supervisor, for his assistance and invaluable support, both technical and moral.

In addition, a vote of thanks goes to Professor Bock Ki Kim for his constant technical advice; to Professor Robert Hollett of the Department of English, who, since my first months of study at Memorial, supervised my efforts in English and now had patiently reviewed the script of this document; to Professor Jader Onofre de Morais, Director of Marine Science Laboratory of the Federal University of Ceará, who provided me with the opportunity to undertake my Master's program; to Professor Antonio Adanto Fonteles-Filho of the Federal University of Ceará, who suggested the collection of data from fish-weirs; to Professor Tereza Cristina Vasconcelos Gesteira who helped me with the classification of species; to Jerome Dominic who prepared the computer programs; and to my parents, brothers and sisters for their motivation during this period out of Brazil.

The author is also very grateful to Professor Jain and Professor Kim for allowing the use of the computer programme which is eventually going to be published in their book.

Financial support came from CIDA (Canadian International Development Agency), without whose aid this study would not have been possible. The manuscript was expertly typed by Mrs. Tina King.

And finally, a special thanks goes to Iracema, my wife, for her love, patience, understanding and support during a very important period in our lives.

TABLE OF CONTENTS

List of Tables

Appendices Tables (Cont'd) Page

List of Figures

CHAPTER I

INTRODUCTION

In the past twenty five years, the concept of species diversity has been much used by ecologists and it suggests something more than the notion of the effective number of species present in a sample (Lloyd, et al., 1968; Hill, 1973). A community (or collection) in which all the individuals belong to one species has no diversity, while one in which every individual belongs to a different species has the maximum possible diversity (Pielou, 1966a). This can be thought as something analogous to statistical variance. In the same way that a variance provides a measurement of the variability of some quantitative variable, a diversity index measures the variability of the species identity. Many indices of diversity have been introduced in ecological literature: for example, the proportion of the i-th species in a community p_i or logarithms of this quantity, Simpson's (1949) index, $1 - \sum p^2_i$ (actually $\sum p^2_i$ in his paper, this modification was proposed by Pielou (1969)); the "information content" as a measure of diversity, (Shannon's (1949) index, $-\sum p_i \log p_i$); the jackknife method proposed by Quenouille (1956) as a means of reducing bias; the Brillouin's (1962) index, useful when a collection is not too large for all its members to be identified and counted (Pielou, 1966); and Pielou's (1966) index, which estimates the community's species diversity based on a sequence of randomly selected sampling units (or quadrats).

Quantitative analysis of the structure of a community, for example, the community of animals living in the area of the fish-weirs at Almofala Beach (Ceará — Brazil), required the measurement of some parameters which can be feasibly tested and inferred. However, in dealing with biological populations, and particularly with fish communities containing many different species, the investigator faces some sampling problems, the greatest of which is the difficulty of collecting necessary information in such a way that it may be deemed to be representative of the population. In this investigation, we have tried to avoid some of these problems, since in our sampling system we used a type of fishing gear which can be considered non-selective in regard to animal size.

There were also other difficulties, for instance, not knowing the true number of species in the community, or how to study the community's structure. These difficulties have led to the development of models that fit species-abundance distributions. In an attempt to find a probability model to fit such data, Fisher et al., (1943) suggested that frequency distributions of species can best be specified by a logarithmic series. Subsequently, many diverse models have been proposed to describe species abundance distributions, but Preston (1948) was the first to use the log-normal distribution to summarize such data.

The purpose of this investigation is to estimate fish diversity utilizing various techniques described in the literature and also an attempt is made to find a suitable model to fit the data of species abundance distributions.

In chapter two we describe the sampling procedures and the fishing gear utilized to collect the necessary data. In an attempt to investigate as many individuals as possible, to minimize the possibility of missed species, the aid of commercial fishery operations was required. The data is summarized in tables and classified into taxonomic levels.

In chapter three the frequency distributions of the species are summarized in species abundance distributions and truncated log-normal distributions are fitted to the data, in order to investigate any consistent interrelation among species caught by different fish-weirs. This procedure is shown, step by step, and illustrated with empirical data.

In chapter four we showed how to estimate the diversity indices proposed by Shannon (1949), Quenouille (1956), and Pielou (1966). Using these indices, three types of parameters were estimated. The first was the community's species diversity index; the second was the species diversity index for each of the sampled fish-weirs; and the third was the community's species diversity index divided into hierarchical components. In some cases, more than one index was used to allow the comparison among the estimators.

The final chapter is devoted to analyzing all estimates and drawing some conclusions about the community under study. It also indicates a possible direction for further research in this area.

CHAPTER II

DATA COLLECTION METHODOLOGY

2.1 Sampling Design

In studying species diversity of marine biological communities, the investigator faces the problem of getting unbiased samples from these communities in such a way that they can be assumed to be representative of each biological population living in an area. In large open sea areas, not feasible for impounding, it is very difficult to obtain such samples, particularly because of the nonrandom individual distributions. Thus, in order to collect the necessary data, the aid of commercial fishery operations is required, in which case, dependence on catch data is obligatory.

In our sampling system, we used a type of gear which can be considered nonselective in regard to animal size, namely the fish-weir, which will be described in the next section. Because of the natural entry of the fish into the weirs with the movement of the tide, the catch species composition may be safely assumed to be representative of that of the community living in the surrounding area.

Pielou (1977) uses the word "community" and suggests that "it means all the organisms in a chosen area that belong to the taxonomic group the ecologist is studying". Thus, the community to be studied in this investigation is comprised of all the animal species that come inshore and make themselves vulnerable to catching by fish-weirs.

A single line of six fish-weirs was chosen for investigation of its commercial catch. This particular line was chosen mainly because of the ease of getting the data when the boats arrived at the beach. Since the gear is lined up from the shore to the sea and since its depth varies with the distance from the coast, each fish-weir can be taken as a stratum covering different depths and distances. Since the animals are collected at least once per day, when the tide is low during the daylight hours, the commercial catch from each weir is considered to be a sampling unit. For each unit, observations both in number and diversity were made for a ten day period in June, 1980. The area covered by this fish-weir line is bounded by the coast and a five meter deep isobath, which is approximately 3,500 meters from shore. Despite the narrowness of this strip, the community of individuals that occur there is assumed to be representative of the species, in both number and diversity, that are found over the whole of the continental shelf.

2.2 Description of the Fishing Gear

The fish-weir is common fishing implement used along the coast of Ceará State, with major concentration at Almofala Beach, in Acaraú County (Ceará - Brazil), latitude 02° 50'S, longitude 40° 09'W (figure 2.1). It is coastal trap-line gear, composed of a guide fence and three compartments, which is arranged is such a manner as to allow the entry of animals deep into the back-end. The fence, the "espia", made of liane netting, runs obliquely to the coast and parallel with the movement of the tide. The animals move with the tide along the espia into the "sala grande", a heart-shaped enclosure, also of liane netting, attached to wooden stakes. The sides of this compartment are rounded inwards and backwards to prevent the animals from escaping once they move in. The second compartment, the "salinha", is a smaller version of the "sala grande", but with a wire netting fence, to enhance the trapping capability of the gear. The third and last compartment, the "chiqueiro" is round-shaped and also made of wire netting fence. This is the most important section because this is where the animals are finally trapped. A scheme of this gear is given in figure 2.2. The size of each compartment and the length of the net of each fish-weir chosen in this investigation are shown in Table 2.1. The table also shows the depth at which the fish were harvested as well as the distance of each weir from the coast. The "espia" size is taken as the total length of the fence; for the other sections, the round measurement of each fence between their tips has been taken. A peculiar aspect of the fish-weir is that the east side of the "sala grande" is shorter than the other side. The traps

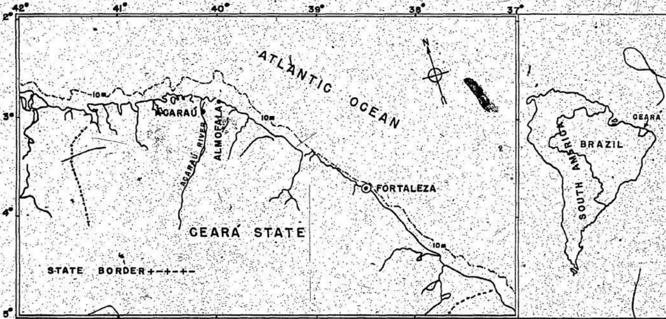


Figure 2.1 - Location of Almofala Beach, in Acaraú County, State of Ceará-Brazil

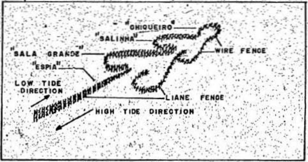


Figure 2.2 - Scheme of a fish-weir used at Almofala (Acaraú-Ceará-Brazil)

TABLE 2.1

Characteristics of the six fish-weirs used at

Almofala Beach (Acaraú - Ceará - Brazil)

Number of fish-weirs	Sizes (meters)						Depth of harvest (meters)	Distance from the coast (km)
	Espia	Sala grande		Salinha	Chiqueiro	Length of the net		
		east fence	west fence					
5	53	33	40	9	16	6.90	4.40	1.3
6	48	33	43	10	17	8.40	4.90	1.5
8	47	27	33	9	17	8.20	4.90	1.9
9	40	13	13	9	16	7.36	4.00	2.1
10	53	16	30	9	15	6.70	4.00	2.3
16	60	13	13	9	15	7.18	4.70	3.5

are imbedded in shallow water areas and set in a straight line to the seaward, with a distance of about 150 meters between each. They are identified by numbers, assigned in crescent order from the coast. The fish-weirs sampled in this investigation were numbers 5, 6, 8, 9, 10 and 16, because they were in operation at the time of this study (figure 2.3).

For the fishing operations, the fishermen use a small sail boat to enter the "chiqueiro", and a small meshed net, which operates as a double stick-net to harvest all the trapped animals. The length of this net, which is made of a vegetable fibre known as "tucum" (Bactris setosa), varies for each fish-weir and its height is proportional to the depth inside the "chiqueiro". It has a 3.5 centimeter mesh size. (figures 2.4 and 2.5).

The importance of the weir from the socio-economic point of view has been discussed by Seraine (1958). Considerations about the fishery production and indices of productivity of different species have been made by Paiva & Nomura (1965), Paiva & Fonteles-Filho (1968), Collyer & Aguiar (1972), and Almeida (1974).

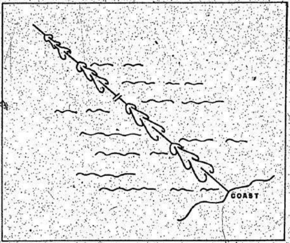


Figure 2.3 - Scheme of a line of fish-weirs used on the coast of the State of Ceará-Brazil.

Figure 2.4 - Fishermen transporting a double stick-net used to harvest the fish inside the "chiqueiro".

Figure 2.5 - Detail of the mesh of the double stick-net used to harvest the fish from the "chiqueiro".

2.3 Data Collected

The estimates of diversity and related statistical analysis in this investigation are based on sixty sampling units, which resulted from observations of the commercial daylight catch of a single line of six fish-weirs, from Almofala Beach (Ceará - Brazil), during ten days in June, 1980. The distribution of the sampling units, per species, and individuals caught per each fish-weir, are shown in Tables A.1 to A.6 (Appendix A).

A total catch of 161,173 individuals caught during this period of study has been classified into three taxonomic levels: 4 classes, 24 families and 44 species. For the sake of simplicity, we have assigned species identification numbers. In the remainder of this investigation, each species will be referred to by its number. These were assigned at the time the species appeared in the catch but they are presented as a natural or phyletic sequence of families, with the species of each family alphabetized by specific names (Table 2.2).

TABLE 2.2

Classification of the species, with respective identification number and quantity of individuals per species, recorded on a line of six fish-weirs, at Almofala Beach (Acaraú-Ceará-Brazil), during ten days of observations in June, 1980.

Class	Family	Species Name	identification number	Common species name	individuals
Crustacea	Palinuridae	Panulirus argus (Latreille)	41	spine lobster	14
		Panulirus laevicauda (Latreille)	42	spine lobster	2
Chondrichthyes	Orectolobidae	Ginglymostoma cirratum (Bonnaterre)	44	nurse shark	1
Osteichtyes	Elopidae	Tarpoon atlanticus (Valenciennes)	20	tarpon	11
	Clupeidae	Opistonema oglinum (Le Sueur)	8	thread herring	110,905
	Engraulidae	Anchoa spinifer (Valenciennes)	21		1
		Anchovia clupeoides (Swainson)	9		35
		Lycengraulis grossidens (Cuvier)	10	snake mouthed sardine	65
	Ariidae	Tachysurus herzbergii (Block)	5		500
		Tachysurus sp.	6		294
		Tachysurus spixii (Agassiz)	7	catfish	32
	Exocoetidae	Hemirhanphus brasiliensis (Linnaeus)	22	ballyhoo	17
	Belonidae	Ablenner hians (Vallenciennes)	23	flat needlefish	222

TABLE 2.2 (Cont'd) page 2

Class	Family	Species Name	identi-fication number	Common species name	individuals
	Pomatomidae	Pomatomus saltatrix (Linnaeus)	24	bluefish	15
	Carangidae	Alectis ciliaris (Block)	14	threadfish	2
		Caranx erysos (Mitchill)	25	Gluerunner hardtail	80
		Caranx hippos (Linnaeus)	3	crevalle jack	176
		Caranx latus (Agassiz)	27	horse-eye jack	1
		Caranx sp.	26		23
		Chloroscombrus chrysurus (Linnaeus)	11	bumper	44,965
		Hemicaranx amblyrhynchus (Cuvier)	12		1
		Oligoplites palombeta (Cuvier)	28		15
		Selene vomer (Linnaeus)	13	lookdown	2,124
		Trachinotus carolinus (Linnaeus)	29	florida pompano	6
		Trachinotus falcatus (Linnaeus)	30	permit	1
	Lutjanidae	Lutjanus jocu (Block & Schneider)	31	dog snapper	1
	Lobotidae	Lobotes surinamensis (Block)	32	tripletail	4
	Pomadasyidae	Genyatremus luteus (Block)	33		196
	Sciaenidae	Cynoscion acoupa (Lacépède)	17	sea trout	5
		Cynoscion leiarchus (Cuvier)	15	sea trout	173
		Cynoscion microlepidotus (Cuvier)	16		6
		Micropogon furnieri (Desmarest)	34		4
	Ephippidae	Chaetodipterus faber (Broussonet)	18	spadefish	45
	Chaetodontidae	Holacanthus ciliaris (Linnaeus)	19	queen anglefish	1

TABLE 2.2 (Cont'd) page 3

Class	Family	Species Name	identi-fication number	Common species name	individuals
	Scaridae	<u>Sparisoma swainson</u>	37		2
	Sphyraenidae	<u>Sphyraena barracuda</u> (Walbaum)	36	great barracuda	58
	Trichiuridae	<u>Trichiurus lepturus</u> (Linnaeus)	4	cutlassfish	371
	Scombridae	<u>Euthynnus alleteratus</u> (Refinesque)	35	little tuna	28
		<u>Scomberonorus cavalla</u> (Cuvier)	2	king mackerel	24
		<u>Scomberomorus maculatus</u> (Mitchill)	1	spanish mackerel	792
	Rehycentridae	<u>Rachycentron canadus</u> (Linnaeus)	40		9
	Ostraciontidae	<u>Lactophrys trigonus</u> (Linnaeus)	38	common trunkfish	2
	Tetradontidae	<u>Lagocephalus laevigatus</u> (Linnaeus)	39	smooth puffer	1
Reptilia	Chaelonidae	<u>Chelonia mydas</u>	43	green turtle	3
					161,173

CHAPTER III

SPECIES ABUNDANCE DISTRIBUTION

3.1 Introduction

The sampling units taken from the community of animals living in the area of the fish-weirs at Almofala Beach (Ceará - Brazil) have exhibited a property common to most ecological communities, in that the inhabitants belong to several species, and the abundance of each species varies greatly (Tables A.1 to A.6). Thus, it will be useful to summarize such data on frequency distribution charts, where the number of species, f(r), containing r representatives (r = 1, 2, ...), are plotted against the number of representatives per species (figures 3.1 to 3.6). Customarily, this frequency distribution represents species-abundance distributions (Poole, 1974; Pielou, 1975 and Slocomb et al., 1977).

In order to investigate any consistent interrelation among species caught by different fish-weirs (strata) it will be required that some form of probability distribution fit those species-abundance distributions with only the parameter values varying from one strata to another, in such a way as to allow the comparison among the values. In an attempt to find a probability distribution to fit such data, Fisher et al., (1943), suggested that those frequency distributions can best be specified by a logarithmic series. Subsequently, many diverse models have been proposed to describe species-abundance distributions, but Preston (1948) was the first to use the log-normal distribution to summarize such data.

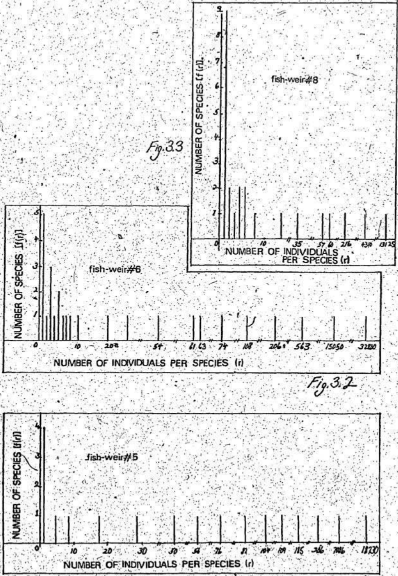


Fig. 3.3

Fig. 3.2

Fig. 3.1

Figures 3.1-3.3 - Frequency distributions of the individuals trapped by fish-weirs numbers 5, 6 and 8, at Almofala Beach (Acaraú-Ceará-Brazil).

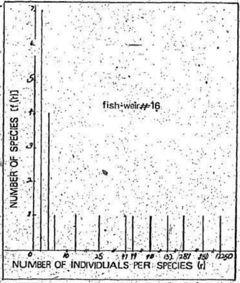


Fig. 3.6

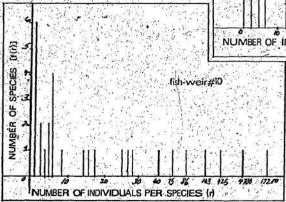
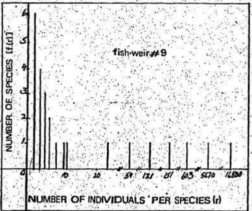


Fig. 3.5

Fig. 3.4



Figures 3.4-3.6 — Frequency distributions of the individuals trapped by fish-weirs numbers 9, 10 and 16, at Almofala Beach (Acaraú-Ceará-Brazil.

In this chapter, species-abundance curves have been fitted to observed frequency distribution using the truncated log-normal distribution.

3.2 Truncated Lognormal Distribution

Preston (1948) found that in a large and diverse community, the frequency distribution of the individuals will follow a normal law after the individuals are grouped on a logarithmic scale. The probability density function for the lognormal distribution is:

$$f(\lambda) = \frac{1}{\lambda\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}\left(\ell n \frac{\lambda}{m}\right)^2\right], \; 0 < \lambda < \infty$$

where λ is the species-abundance. Let $r = \ell n\ \lambda$, then $\frac{d\lambda}{dr} = \lambda$, so that

$$f(r) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(r - \ell n\ m)^2\right].$$

Now, make $\alpha = \ell n\ m$, thus,

$$f(r) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(r - \alpha)^2\right], \; -\infty < r < \infty,$$

where σ is the logarithmic standard deviation,

α is the position of the mode, and

r is the position of an observed number of species.

It is both natural and convenient to plot r the number of individuals belonging to a given species on a logarithmic scale, since an observed species abundance histogram is usually L-shaped (figures 3.1 to 3.6), with very few high frequencies for low values of r and a long tail representing the few abundant species.

It is generally agreed that in ecological work, and particularly with fishery data, it is very difficult to ensure that the sample collected represents the entire community. Pielou (1977) remarked that there are some species so rare that they are not expected to be found in a sample of the size at hand. Thus, the truncation of the curve on the left is inevitable. Consequently, the number of species in the community, s, is unknown and must be estimated. Preston called this point of truncation the "veil line". If every species in the community has been sampled, no difficulty arises in the estimation of the location and scale paremeters. Since this is not the case, the fitting of a lognormal distribution will be described "in recipe form" (Pielou 1975) and illustrated with field data from fish-weir number 10.

The procedure consists of converting the observed variate, r, to logarithmics of base 10 and then fitting a normal distribution to the r's. Since we are treating r as a continuous variate, we should substitute the discrete values 0, 1, 2, ..., r, ..., for the intervals $(0, \frac{1}{2}]$, $(\frac{1}{2}, 1\frac{1}{2}]$, $(1\frac{1}{2}, 2\frac{1}{2}]$, ..., $(r - \frac{1}{2}, r + \frac{1}{2}]$... . Since the distribution is zero-truncated and the "empty" (unrepresented) species are unobservable, the value $r = 0$ is missing from the discrete data, and the interval $(0, \frac{1}{2}]$ is missing from its continuous representation. Hence the normal distribution to be fitted to the r's is truncated on the left at $r_o = \log 0.5$.

The procedure, step by step, and the numerical results (Table 3.1) are as follows:

1) make $x^* = \log r + \frac{1}{2}$;

2) obtain the mean of the number of individuals in each species sample ($\bar{x}^*$) and the sampled variance (V^2) given by.

$$\bar{x}^* = \frac{\sum_r x^* f(r)}{s}$$

and

$$V = \frac{\sum_r (x^* - \bar{x}^*)^2 f(r)}{s}$$

where s is the total number of species in the sample;

3) calculate

$$\gamma = \frac{V^2}{(\bar{x}^* - r_o)}$$

where $r_o = \log 0.5 = -0.30103$;

4) from Table 1 in Cohen (1961) obtain the "auxiliary estimation function" $\hat{\theta}$ corresponding to this γ;

5) obtain estimates $\hat{\mu}$ and $\hat{\sigma}^2$ of the mean and variance of r, from

$$\hat{\mu} = \bar{x}^* - \hat{\theta}\ (\bar{x}^* - r_o)$$

and

$$\hat{\sigma}^2 = V^2 + \hat{\theta}\ (\bar{x}^* - r_o)^2;$$

6) obtain the standardized normal variate, say Z_o, corresponding to the truncation point r_o by making

$$Z_o = \frac{r_o - \hat{\mu}}{\hat{\sigma}};$$

7) from the standard normal tables, find $p_o = p_r\ (Z \leq Z_o)$, the area under the normal curve to the left of Z_o;

8) hence, obtain

$$\hat{S}^* = \frac{s}{1 - p_o},$$

the estimated number of species in the community;

9) compile Table 3.1 as indicated - the Roman numerals refer to the columns in the table which show:

(i) the value of r,

(ii) $\log\ (r + \frac{1}{2})$ - the upper boundary of each class-interval,

(iii) the number of species, that is, the observed frequencies f(r),

(iv) and (v) computations used to obtain the mean and variance of the samples,

(vi) the variate value in (ii) in standardized form - thus (vi) $= \frac{x^* - \hat{\mu}}{\sigma}$,

(vii) the area under the normal curve,

(viii) the accumulated expected frequencies,

(ix) differences between successive entries in (viii).

Hence these are the desired expected frequencies (f_e) of the (iii);

10) judge the goodness of fit of what is expected to the observed frequencies, using an χ^2 test. The number of degrees of freedom (d.f.) is three less than the number of frequencies compared, since two d.f. were lost by using the estimates $\hat{\mu}$ and $\hat{\sigma}$, and one additional was lost because of the fixed total number of species.

TABLE 3.1

The fitting of a truncated lognormal distribution and estimation of the needed parameters to the data of fish-weir number 10.

i	ii	iii	iv	v	vi	vii	viii	ix		
r	$\log(r + \frac{1}{2}) = x^*$	f(r)	$x^*f(r)$	$(x^* - \bar{x}^*)^2\ f(r)$	$(x^* - \hat{u})/\sigma = z$	$\Phi(z)$	$\hat{s}^*\ \Phi(z)$	fe	goodness-of-fit	
									fe	f(r)
0	-0.30103	0	0	0	-1.41110	0.0793	2.50	(2.50)		
1	0.17609	6	1.05654	2.78649	-0.75248	0.2266	7.14	4.64	7.78	8
2	0.39794	2	0.79588	0.42252	-0.44623	0.3264	10.28	3.14		
2 ┤ 4	0.65321	3	1.95963	0.12529	-0.09385	0.4641	14.62	4.34	9.06	8
4 ┤ 8	0.92942	5	4.64710	0.02581	0.28743	0.6141	19.34	4.72		
8 ┤ 16	1.21748	3	3.65244	0.38860	0.68507	0.7549	23.78	4.44	7.82	6
16 ┤ 32	1.51188	3	1.80956	1.28436	1.09147	0.8621	27.16	3.38		
32 ┤ 64	1.80956	1	6.32670	0.90628	1.50239	0.9332	29.40	2.24		
64 ┤ 128	2.10890	3	-	4.69748	1.91560	0.9726	30.64	1.24	4.34	7
128 ┤ 256	2.40909	-	-	-	2.32999	0.9901	31.19	0.55		
> 256	∞	3		-	-	1.0000	31.50	0.31		

Estimated parameters based on data of the fish-weir number 10. The computations are led in Table 3.1

$s = 29$

$\bar{x}^* = \frac{\Sigma x^* f(r)}{s} = 0.85757$

$V^2 = \frac{\Sigma(x^* - \bar{x}^*)^2 \cdot f(r)}{s} = 0.36679$

$\gamma = \frac{V^2}{(\bar{x}^*) - x_o} = 0.31658$

$\hat{\theta} = 0.1177$

$\hat{u} = 0.72120$

$\hat{\sigma}^2 = 0.52479$

$\hat{\sigma} = 0.72442$

$Z_o = -1.4111$

$p_o = Pr(Z \leq Z_o) = Pr(Z \leq -1.4111) = 0.0793$

$\hat{S}^* = \frac{s}{1 - p_o} = 31.5$

$\chi^2 = 2.17 < \chi^2_{.05,1} = 3.841$

The value of chi-square clearly shows that truncated lognormal distribution fit to the data of fish-weir number 10. For the community of animals in the area of fish-weirs at Almofala Beach, the goodness of fit of the truncated lognormal distributions are shown in Appendix B, Table B.1. Further analysis of the lognormal distribution is given in the last chapter.

CHAPTER IV

DIVERSITY INDICES CONSIDERED

4.1 Introduction

The models that have been devised to fit the species abundance distributions can often be based on contrasting sets of initial premises; therefore, the same predictions can sometimes be yielded by two or more contrasting models. Thus, in this chapter, we will search for some form of descriptive statistics which can be used even when no theoretical model can be found to fit the data.

As noted in the last chapter, the community of individuals under investigation has exhibited a property in which both the number of species and the relative proportion of each play important roles. This property has led to the development of single statistics which are known as "indices of diversity". Three conditions, which we will describe, are desirable in such an index (Pielou 1977). Let us suppose we are dealing with a community that can be classified into s species. Every individual belongs to one and only one species, and the probability that a randomly selected individual will belong to the species s_i is p_i. Thus,

$$\sum_{i=1}^{s} p_i = 1.$$

As a measure of the diversity of the community, we wish to find a function of p_i, $H'(p_1, p_2, \ldots, p_s)$ say, that meets the following conditions:

Condition 1. For a given s, the function takes its greatest value when

$$p_i = \frac{1}{s} \text{ for all } i.$$

Denoting this greatest value by L(s),

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

Such a community is said to be completely even.

Condition 2. Given two completely even communities, one with s species and another with s + 1, the latter should have the greater diversity.

Condition 3. Suppose the community individuals are subject to two separate classifications (not necessarily independent). Assume an A-classification with a classes, and a B-classification with b classes.

Let p_j be the probability that a randomly selected individual will belong to class A_j. Thus,

$$\sum_{j=1}^{a} p_j = 1.$$

Let q_k be the probability that a randomly selected individual will belong to class B_k. Thus,

$$\sum_{k=1}^{b} q_k = 1.$$

Then, the double classification yields a x b different classes, $A_j\ B_k$ (j = 1, 2, ..., a; k = 1, 2, ..., b), and the probability that a randomly selected individual will belong to the class $A_j\ B_k$ may be written π_{jk}. Clearly, if the A- and B-classifications are independent,

$$\pi_{jk} = p_j\ q_k.$$

Suppose that A- and B-classifications are mutually dependent, then

$$\pi_{kj} = p_j\ q_{jk},$$

where q_{jk} is the conditional probability that an individual will belong to B_k, given that it belongs to A_j.

For the diversity of the doubly classified community we may write

$$H'(AB) = H'(\pi_{11}, \pi_{12}, \ldots, \pi_{ab}).$$

For the diversity under the B-classification within the class A_j, write

$$H'_j(B) = H'(q_{j1}, q_{j2}, \ldots, q_{jb}).$$

And now make

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

for the mean diversity index under the B-classification within all the A- classes.

Hence, we get

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

Independence of A- and B-classifications implies $q_{jk} = q_k$ for all j. Therefore, we have

$$H'(AB) = H'(A) + H'(B);$$

This condition is needed, for instance, when the individuals in a community are classified into taxonomic levels.

In an ecological context, condition (1) should have the following interpretation: for a community with a given number of species, the measure of diversity will be maximum when all the species are present in equal proportions. The usefulness of condition (3) will be discussed in Section 4.5

Based on these conditions, several indices of diversity have been introduced in ecological literature and three of them are applied in this investigation: the Shannon-Wiener, or simply Shannon (1949); the Pielou's sequencial estimate (1966); and the so-called "Jackknife method", proposed by Quenouille (1956). These indices will be described in the following sections.

Using these indices, and the data in Tables A.1 to A.6, three types of parameters were estimated. The first was the community's species diversity index; the second was the species diversity index for each of the sampled fish-weirs (stratum); and the third was the community's species diversity index divided into hierarchical components.

The community's species diversity was estimated using both Pielou's (1966), and Shannon's (1949) indices. The Jackknife method on Shannon's index was also considered. The diversity index for each weir was estimated using both Shannon's (1949), and the Jackknife method. This parameter was estimated in order to investigate the special differences in the diversity index in the area of the fish-weir. The hierarchical diversity was measured by Shannon's (1949) index, after the community was classified into three taxonomic levels, by class, family and species.

4.2 Shannon-Wiener Index of Diversity

Having specified the three conditions that H' is to satisfy, Khinchin (1957) and Pielou (1969 and 1977) have shown that the only function of the p_i values having these three conditions is:

$$H'(p_1, p_2, \ldots, p_s) = -C \sum_{i=1}^{s} p_i \log p_i$$

where C is a positive constant;

p_i is the proportion of the community belonging to the i^{th} species.

Making C = 1, we may, therefore, take as an index of diversity,

$$H' = -\sum_{i=1}^{s} p_i \log p_i \qquad (4.1)$$

This was originally proposed by Shannon (Shannon and Weaver, 1949).

Goldman (1953) has emphasized that Shannon is index H' is defined to estimate the average diversity from a sample, when the community is large enough for all its members to be identified and counted.

An estimate $\hat{H}'$ of H' may be given by

$$\hat{H}' = -\sum_{i=1}^{s} \hat{p}_i \log \hat{p}_i, \qquad (4.2)$$

where $\hat{p}_i$ is the proportion of the ith species in the sample.

Basharin (1959) has shown that $\hat{H}'$ is a biased estimation of H'. The magnitude of this bias depends on how close the number of species present in the sample is to the true number of species (s) in the community. He has also shown that $\hat{H}'$ is considtence and asymptotically normal.

An estimator of its variance in large samples is given by

$$\text{var}\ (\hat{H}') = \frac{1}{N}\ [\sum \hat{p}_i (\ln \hat{p}_i)^2 - \hat{H}'^2], \tag{4.3}$$

where ℓn is the natural logarithms.

Even though we do not know s, the total number of species that may occur in this community, the bias that eventually may occur in estimating H', can be disregarded if we take into consideration the type of fishing gear used for sampling this community (Section 2.2), and the size of the sample analyzed.

Thus, using equation 4.2, two types of parameters were estimated. The first, the community's species diversity, was estimated after the totals per species of each fish-weir (Tables A.1 to A.6) had been gathered. Whereas, for the second, the equation 4.2 was applied for each total in table A.1 to A.6, to estimate the species diversity index for each fish-weir.

To estimate the variance of these parameters, equation 4.3 was used with the above procdure. The numerical results will be discussed in the next chapter.

4.3 Pielou's Sequencial Estimate

It is known that $\hat{H}'$ is a biased estimator of H' and that the use of the former requires the assumption that the sampling units (s.u.'s) examined constitute a random selection from the community whose diversity is being estimated. But in practice, and particularly with the kind of community being investigated, it is very difficult (perhaps impossible) to ensure a completely random selection. To deal with this difficulty, Pielou (1966) has proposed a method which estimates not only H' but also its standard error, from a series of randomly selected units.

Recall that from each one of the six fish-weirs observed for a ten day period a total of $n = 60$ sampling units was yielded. These are now to be taken in random order and added, one after another, to a growing pool of s.u.'s.

Let N_{xi} be the number of individuals of the ith species in the xth s.u.

Let $M_{ki} = \sum_{x=1}^{k} N_{xi}$ be the number of individuals of the ith species in the pool of the first k s.u.'s.

Let $M_k = \sum_{i=1}^{s} M_{ki}$ be the number of individuals of all species in this pool.

Also, let

$$H_k = \frac{1}{M_k} \ln \frac{M_k!}{\prod_{i=1}^{s} M_{ki}!} \qquad (4.4)$$

be the diversity, as measured by the Brillouin's index. For our data, the values of M_k and H_k are shown in Table 4.1.

For a general estimation problem, Brillouin's (1962) index should be used when a collection is not loo large for all its members to be identified and counted. It is defined as

$$H = \frac{1}{N} \log \frac{N!}{N_1! \, N_2! \ldots N_s!}$$

where N is the total number of individuals, s the number of species, and N_i the number of individuals in the ith species, so that $\sum_{i=1}^{s} N_i = N$ and log means logarithm of base 10. The index H gives the diversity per individual.

Now, returning to expression (4.4), if H_k is plotted against k, and if n is large enough, then it will be found that H_k increases (not necessarily monotonically) at first, and then levels off when the pool of s.u.'s has become large enough for its contents to provide an adequate representation of the community as a whole. (See figure 4.1).

Suppose the curve shows no upward trend, when k > t. Let

$$h_k = \frac{M_k H_k - M_{k-1} H_{k-1}}{M_k - M_{k-1}} \qquad (4.5)$$

and compute h_k where $k = t + 1, t + 2, \ldots, n$. In our case we choose the value of $t = 54$, because at this point, figure 4.1 shows the curve

TABLE 4.1

Values of the H_k, computed for the growing pool of randomly ordered sampling units (M_k), taken from the community of animals living in the area of the fish-weirs at Almofala Beach, (Ceará - Brazil).

Sampling units(k)	H_k	M_k
1	0.0183	3,509
2	0.2993	10,338
3	0.3502	14,708
4	0.4380	20,720
5	0.4935	25,746
6	0.5568	30,713
7	0.5641	35,948
8	0.5898	40,516
9	0.6071	43,411
10	0.6211	46,076
11	0.6435	59,945
12	0.6537	61,931
13	0.6522	65,877
14	0.6689	66,148
15	0.6655	68,759
16	0.6685	70,116
17	0.6543	74,117
18	0.6564	77,130

Table 4.1 (cont'd)

Sampling units(k)	H_k	M_k
19	0.6712	87,690
20	0.6858	88,001
21	0.6893	88,658
22	0.6966	89,114
23	0.6993	90,356
24	0.7033	90,659
25	0.7051	92,460
26	0.7060	93,812
27	0.7076	96,428
28	0.7134	96,665
29	0.7177	99,767
30	0.7388	115,279
31	0.7367	119,356
32	0.7441	120.024
33	0.7529	120,209
34	0.7549	120,516
35	0.7516	122,825
36	0.7532	122,903
37	0.7532	126,931
38	0.7560	128,051
39	0.7536	130,214

Table 4.1 (cont'd)

Sampling units(k)	H_k	M_k
40	0.7553	131,425
41	0.7578	133,470
42	0.7635	137,519
43	0.7651	141,545
44	0.7631	142,658
45	0.7639	142,676
46	0.7642	143,223
47	0.7698	145,301
48	0.7675	146,558
49	0.7689	150,099
50	0.7718	150,223
51	0.7752	150,671
52	0.7788	151,160
53	0.7808	151,731
54	0.7838	152,514
55	0.7879	153,442
56	0.7880	154,683
57	0.7940	156,063
58	0.7974	157,413
59	0.7979	158,963
60	0.7963	161,173

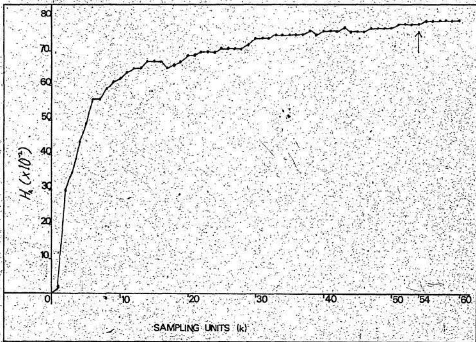


Figure 4.1 – Plots of H_k versus k. Values from Table 4.1

to be leveling off. The values of h_k are given in Table 4.2.

Table 4.2

Values of h_k where k = 55, 56, ..., 60.

k	h_k
55	1.48690
56	0.790394
57	1.469701
58	1.191897
59	0.844476
60	0.687698

It can be shown that an estimate of H' is given by

$$\bar{H}' = \frac{1}{n - t} \sum_{k = t + 1}^{n} h_k = \bar{h}; \qquad (4.6)$$

the sampling variance of this estimate is given by

$$\text{var } (\bar{H}') = \text{var } (\bar{h}) = \frac{1}{n(n - 1)} \left(\sum_{k = t + 1}^{n} h_k^2 - n \bar{h}^2 \right). \qquad (4.7)$$

The numerical results will be discussed in the next chapter.

4.4 Jackknife Estimate of Diversity

The following jackknife technique was introduced by Quenouille (1956) as a means for reducing the bias of an estimator.

Let $y_1, y_2, \ldots, y_N$ be N independent observations with distribution depending on a parameter θ. Divide the observations into g groups of k observations (sampling units) with $N = gk$. Let $\hat{\theta}_g$ be an estimate of θ based on all N observations and $\hat{\theta}^k_{g-1}$ be the sub-estimate obtained after deleting the kth group of observations. If θ^*_k, called a pseudovalue, is defined as

$$\theta^*_k = g\,\hat{\theta}_g - (g-1)\,\hat{\theta}^k_{g-1}$$

then the jackknife estimate of θ is

$$\hat{\theta}_J = \frac{1}{g} \sum_{k=1}^{g} \theta^*_k$$

and an estimate of the variance of $\hat{\theta}_g$ is

$$\text{var}\,(\hat{\theta}_g) = \frac{\sum_{i=1}^{g} (\theta^*_k - \hat{\theta}_J)^2}{N(N-1)} .$$

Now consider Shannon's (1949) index in estimating the diversity per fish-weir.

Let p_i be the proportion of the ith species in the kth s.u.; $1 \leq i \leq s$ and $1 \leq k \leq g$, where s may vary for each s.u., and $g = 10$.

Let $\hat{H}'_J$ be an estimate of H' based on all s.u.'s, then

$$\hat{H}'_J = -\sum_{i=1}^{s} \sum_{k=1}^{g} p_{ik} \ln p_{ik}.$$

For the estimate with the kth s.u. omitted, let k' have the same range as the index k and let

$$\hat{H}'^{(-k)}_J = -\sum_{i} \sum^{k} p_{ik'} \ln p_{ij'},$$

where $\sum^{k}$ indicates summation over the range of k' but not k.

So the pseudovalues can be defined as

$$\hat{H}'^{*}_k = g\,\hat{H}'_J - (g-1)\,\hat{H}'^{(-k)}_J.$$

The jackknife estimate is the average of these pseudovalues; so

$$\hat{H}'_{Jk} = \frac{1}{g} \sum_{k=1}^{g} \hat{H}'^{*}_k$$

and its estimated variance is

$$\mathrm{Var}\,(\hat{H}'_J) = \frac{1}{g(g-1)} \sum_{k=1}^{g} (\hat{H}'^{*}_k - \hat{H}'_{Jk}).$$

For the estimate of the community's species diversity using Shannon's (1949) index, the procedure is the same, except that we should make $g = 60$. These estimates will be discussed in the next chapter. The numerical results are shown in Appendix C.

4.5 Hierarchical Diversity

The diversity indices considered in the previous sections take no account of the hierarchical nature of biological classification. Since the community being investigated was partitioned hierarchically into class, family, and species, the application of condition 3, (Section 4.1), is straightforward. The Shannon (1949) index ($\hat{H}'$) has been used to measure the community's total diversity, as well as its components - the class diversity, the family diversity within each class, and the species diversity within each separate family and class.

Suppose that there are c classes and that the number of individuals in the ith class is N_i ($i = 1, \ldots, c$; $\sum_{i=1}^{c} N_i = N$); that there are f_i families in the ith class and N_{ij} individuals in the jth family of the ith class ($j = 1, \ldots, f_i$; $\sum_{j=1}^{f_i} N_{ij} = N_i$); and that there are s_j species in the jth family and K_{ijk} individuals in the kth species of the jth family of the ith class ($k = 1, \ldots, s_j$; $\sum_{k=1}^{s_j} N_{ijk} = N_{ij}$). The following notations are used to denote total diversity, class diversity, and family diversity, etc.:

$\hat{H}'(SFC) \equiv$ the species diversity of the community - that is, the total diversity;

$\hat{H}'(C) \equiv$ the class diversity of the whole community;

$\hat{H}'_i(F) \equiv$ the family diversity within the ith class;

$\hat{H}'_C(F) = \sum_{i=1}^{c} \frac{N_i}{N} \hat{H}'_i(F)$ is the weighted mean of the family diversity in all c classes;

$\hat{H}_{ij}(S) \equiv$ the species diversity within the jth family of the ith class;

$\hat{H}'_{CF}(S) = \sum_{i=1}^{c} \sum_{j=1}^{f_i} \frac{N_{ij}}{N} \hat{H}'_{ij}(S)$ is the weighted mean of the species diversity within the family in all c classes.

For a triply classified community the relationship is:

$$\hat{H}'(SFC) = \hat{H}'(C) + \hat{H}'_C(F) + \hat{H}_{CF}(S).$$

These estimates will be discussed in the next chapter.

CHAPTER V

ANALYSIS AND CONCLUSIONS

It was found that the lognormal model fitted to species abundance distributions (figures 3.1 to 3.6), in some cases, showed discrepancies, verified by using the χ^2-test for goodness-of-fit, in the observed data. The following are the needed parameters, estimated to fit a truncated lognormal distribution to the data of the community of animals living in the area of the fish-weirs at Almofala Beach, as well as the χ^2-test for goodness-of-fit, which in thie case showed a good fit. The computations are shown in Appendix B, Table B.1.

$\bar{x}^* = 1.09563$	$\hat{u} = 0.54521$	$Z_o = -0.7174$
$V^2 = 0.62285$	$\hat{\sigma}^2 = 0.39160$	$p_o = 0.2358$
$\gamma = 0.44596$	$\hat{\sigma} = 1.17966$	$\hat{S}^* = 57.58 \approx 58$
$\theta = 0.3941$	$\chi^2 = 2.879 \quad (P < 0.05)$	

The figures 3.1 to 3.6 depicits the consistency in the shape of the species abundance distributions. This is worth considering because it may be a contribution for further research in this area.

The estimates of the community's species diversity, arrived at by using Pielou's ($\bar{H}'$), Shannon's ($\hat{H}'$) and Jackknife ($\hat{H}_{Jk}$) indices, shows very similar estimates for the last two indices and much smaller variance than Pielou's, whereas, Shannon's estimate shows the smallest variance (Table 5.1).

TABLE 5.1

Community's species diversity indices, sampled variances ($\hat{\sigma}^2$) and 95% confidence interval, estimated for $\tilde{H}'$, $\hat{H}'$ and $\hat{H}_{Jk}$ indices.

Indices of diversity	$\hat{\sigma}^2$	95% confidence interval
$\tilde{H}' = 1.078526$	0.124652	$0.386526 < \tilde{H}' < 1.770525$
$\hat{H}' = 0.797048$	0.000006	$0.792247 < \hat{H}' < 0.801849$
$\hat{H}_{JK} = 0.798567$	0.000947	$0.738261 < \hat{H}_{Jk} < 0.858874$

The species diversity index, for each fish-weir, estimated by using both the Shannon and Jackknife indices, shows, again, similar estimates, whereas, the Shannon index gives consistently much smaller variance (Table 5.2).

The following t-test, proposed by Bowman et al., (1969), was used to see if there is a statistical difference in species diversity (estimated by Shannon's index) of fish-weirs numbers 8 and 16, (since they showed the smallest and the biggest diversity, respectively). Let

$$t = \frac{\hat{H}_8' - \hat{H}_{16}'}{[\mathrm{var}(\hat{H}_8') + \mathrm{var}(\hat{H}_{16}')]^{1/2}}$$

The null hypothesis is H_o: $\hat{H}_8' = \hat{H}_{16}'$. The degrees of freedom of the test is

TABLE 5.2

Species diversity indices and respective sampled variances, estimated for each fish-weir, using both Shannon's and jackknife indices.

Fish-weir number	Shannon's index		Jackknife index	
	$\hat{H}'$	$\hat{\sigma}^2_{\hat{H}'}$	$\hat{H}_{Jk}$	$\hat{\sigma}^2_{H_{Jk}}$
5	0.808076	0.000038	0.787298	0.007987
6	0.771433	0.000015	0.789816	0.003674
8	0.697978	0.000045	0.716237	0.010842
9	0.788565	0.000042	0.786590	0.008588
10	0.747718	0.000052	0.746312	0.007556
16	0.843692	0.000032	0.847534	0.002409

$$df = \frac{[\mathrm{var}(\hat{H}'_8) + \mathrm{var}(\hat{H}'_{16})]^2}{\frac{\mathrm{var}(\hat{H}'_8)^2}{N_8} + \frac{\mathrm{var}(\hat{H}'_{16})^2}{N_{16}}},$$

where N_8 and N_{16} are the number of individuals caught by fish-weirs 8 and 16, respectively. The value of $t = 16.5236$ with 36,590 degrees of freedom exceeds the 5% probability level, thus showing a significant difference between estimated diversity indices of the two weirs.

The plotting of $\hat{H}'$ against the distance from the coast and the harvesting depth of the "chiqueiro", (figure 5.1), suggests that high depth implies a low diversity index and vice versa. The former result and this graident suggests a spatial difference in diversity indices between weirs located near the coast and those located farther out.

The hierarchical diversity, measured by Shannon's index, gives the following results:

$\hat{H}'(C) = 0.001316$; $\hat{H}'_C(F) = 0.721948$; $\hat{H}'_{CF}(S) = 0.073758$.

Hence,

$$\hat{H}'(SFC) = 0.797002.$$

It seems that it will be worthwhile to continue collecting basic data for this investigation for a whole year, or perhaps several years, and see how the parameters of the community behave with respect to a fitted species abundance model and diversity indices, when samples are progressively enlarged.

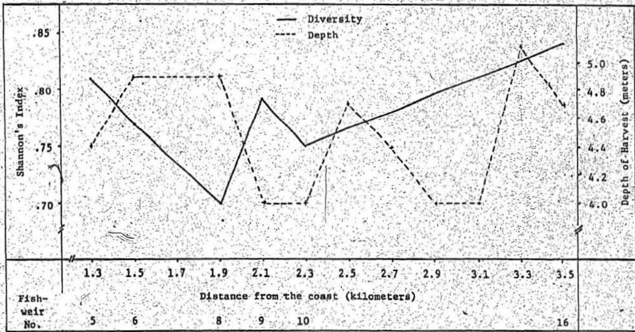


Figure 5.1 - Plots of the estimated Shannon's index of diversity versus distance from the coast and the harvesting depth of the "chiqueiro"

REFERENCES

Almeida, H.T. - 1974 - Sobre a producão pesqueira de alguns currais-de-pesca do Ceará - Dados de 1971 a 1973. Bol. Ciên. Mar, Fortaleza, (26): 1 - 9, 2 figures.

Bailey, R.M. - 1970 - A list of common and scientific names of fishes from the United States and Canada. American Fisheries Society, Special Publication, No. 6, 3rd ed., Washington.

Basharin, G.P. - 1959 - Theory Probably Application, 4, 333 - 336.

Bowman, K.O., Hutcheson, K., Odum, E.P., & Henton, L.R. - 1969 - International Symposium on Statistical Ecology. Vol. 3. University Park: Pennsylvania State University Press.

Brillouin. L. - 1962 - Science and Information Theory, 2nd edition. Academic Press, New York.

Cohen, A.C. Jr. - 1961 - Tables for maximum likelihood estimates: singly truncated and singly censored samples. Technometrics, 3: 535 - 541.

Collyer, C.E. & Aguiar, A.D. - 1972 - Sobre a producão pesqueira de alguns currais-de-pesca do Ceará - Dados de 1968 a 1970. Bol. Ciên. Mar, Fortaleza, (24): 1 - 9, 2 figures.

Fisher, R.A., A.S. Corbet & C.B. Williams - 1943 - The relation between the number of species and the number of individuals in a random sample of an animal population. J. Anim. Ecol. 12: 42 - 58.

Goldman, S - 1953 - Some fundamentals of information theory. In "Information Theory in Biology" (H. Quastler, Ed.). University of Illinois Press, Urbana.

Hill, M.O. - 1973 - Diversity and evenness: A unifying notation and its consequences. Ecology, 54 (2): 427 - 432.

Khinchin, A.I. - 1957 - Mathematical Foundations of Information Theory. Dover, New York.

Lloyd, M., R.F. Inger & F.W. King - 1968 - On the diversity of reptile and amphibian species in a bornean rain forest. The American Naturalist, 102 (928): 497 - 515.

Molles, Jr., M.C. - 1978 - Fish species diversity on model and natural reef patches: Experimental insular biogeography. Ecological Monographs, 48: 289 - 305.

Paiva, M.P. & Fonteles-Filho, A.A. - 1968 - Sobre a producão pesqueira de alguns currais-de-pesca do Ceará - Dados de 1965 - 1967. Bol. Est. Biol. Mar. Univ. Fed. Ceará, Fortaleza (16): 1 - 8.

Paiva, M.P. & Nomura, H. - 1965 - Sobre a producão pesqueira de alguns currais-de-pesca do Ceará - Dados de 1962 a 1964. Arq. Est. Biol. Mar. Univ. Ceará, Fortaleza, 5 (2): 175 - 214, 42 figures.

Pielou, E.C. - 1966 - The measurement of diversity in different types of biological collections. J. Theoret. Biol., 13: 131 - 144.

Pielou, E.C. - 1966a - Species-diversity and pattern-diversity in the study of ecological succession. J. Theoret. Biol. 10: 376 - 383.

__________ - 1969 - An Introduction to Mathematical Ecology. Wiley-Interscience, New York. 286 pp.

__________ - 1975 - Ecological Diversity. Wiley, New York.

__________ - 1977 - Mathematical Ecology. Wiley, New York.

Poole, R.W. - 1974 - An Introduction to Quantitative Ecology. McGraw-Hill, New York.

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

Quenouille, M. - 1956 - Notes on bias in estimation. Biometrika, 43: 353 - 360.

Seraine, F. - 1958 - Curral-de-pasca no litoral cearense. Bol. Antropologia, Fortaleza, 2 (1): 21 - 44, 12 figures.

Shannon, C.E. & W. Weaver - 1949 - The Mathematical Theory of Communication. University of Illinois Press, Urbana.

Simpson, E.G. - 1949 - Measurement of diversity. Nature, 163: 688.

Slocomb, J., B. Stauffer & K.L. Dickson - 1977 - On fitting the truncated lognormal distribution to species-abundance data using maximum likelihood estimation. Ecology, 58: 693 - 696.

Zahl, S. - 1977 - Jackknifing an index of diversity. Ecology, 58: (4): 907 - 913, 2 figures.

APPENDIX A

Basic Data: Distributions of the Sampling Units

TABLE A.1

Distribution of the sampling units, per species and individuals, trapped in fish-weir number 5, during ten days in June, 1980

Species number	Sampling Units (*)										Total
	17	18	19	21	22	23	24	25	26	27	
1	8	3	2	70	1	8	3	—	4	10	109
2	2	3	—	2	—	—	—	—	1	—	8
4	11	20	9	—	9	—	4	12	6	11	82
5	15	84	21	—	6	—	8	90	30	52	306
6	—	—	—	16	16	3	3	7	10	1	56
8	80	100	50	9,200	3,000	2,000	2,000	300	1,000	1,000	18,730
10	20	30	—	—	—	—	—	—	—	—	50
11	15	—	—	4,580	1,500	600	70	20	200	100	7,085
13	50	15	—	—	30	—	50	20	—	20	185
15	37	16	14	—	—	—	—	3	4	2	76
18	3	—	3	—	2	—	15	—	2	4	29
24	—	—	—	—	1	—	—	—	—	—	1
28	—	—	3	—	—	—	1	—	—	—	4
29	—	—	—	1	—	—	—	—	—	—	1
32	—	—	1	—	—	—	—	—	—	—	1
33	30	40	8	—	3	—	7	1	4	11	104
36	—	—	13	—	—	—	1	3	—	—	17
40	—	—	—	—	—	—	1	—	—	—	1

(*) Days in which the commercial catch was observed.

TABLE A.2

Distribution of the sampling units, per species and individuals, trapped in fish-weir number 6, during ten days in June, 1980

Species number	Sampling units (*)										Total
	17	18	20	21	22	23	24	25	26	27	
1	10	4	13	7	15	17	18	4	8	12	108
2	2	—	2	—	—	—	1	—	—	—	5
3	—	—	1	5	54	—	2	—	1	—	63
4	6	15	8	—	—	4	2	—	20	6	61
5	—	1	—	—	2	1	—	50	—	—	54
6	—	—	—	—	21	6	41	7	131	—	206
8	400	—	2,500	9,000	6,900	4,000	2,000	2,000	—	6,000	32,800
11	200	50	1,500	6,000	3,500	1,200	800	1,000	—	800	15,050
13	30	—	—	500	2	—	—	30	—	1	563
15	1	2	1	—	1	3	6	4	6	—	24
16	—	—	—	—	—	—	—	6	—	—	6
17	—	—	—	—	4	—	—	—	1	—	5
18	—	—	—	—	5	—	5	—	—	—	10
20	1	—	—	—	—	—	—	—	—	—	1
21	1	—	—	—	—	—	—	—	—	—	1
22	3	—	—	—	—	—	—	—	—	—	3
23	—	—	—	—	—	—	—	—	10	10	20
24	—	—	—	—	2	3	—	1	1	—	7
28	1	—	—	—	—	—	—	—	—	—	1
29	—	—	—	—	1	1	—	—	—	—	2
32	—	—	1	—	—	—	—	—	—	—	1
33	1	—	1	—	51	—	15	—	6	—	74
34	—	—	1	—	1	—	—	—	1	—	3
35	—	—	—	—	—	—	3	—	—	—	3
36	1	6	—	—	—	—	1	—	—	—	8
40	—	—	—	—	1	—	3	—	—	—	4
43	—	—	—	—	—	—	1	—	—	—	1

(*) Days in which the commercial catch was observed.

TABLE A.3

Distribution of the sampling units, per species and individuals, trapped in fish-weir number 8, during ten days in June, 1980

Species Number	Sampling units (*)										Total
	18	19	20	21	23	24	25	26	27	28	
1	8	4	–	12	7	4	4	5	7	6	57
2	–	2	–	2	–	1	–	–	–	–	5
3	–	–	2	1	–	4	–	–	–	–	7
4	15	6	12	7	–	4	4	3	2	7	60
5	–	–	–	–	–	–	1	–	–	–	1
6	–	–	2	–	–	–	–	–	–	–	2
8	75	150	900	2,200	2,000	–	1,800	1,500	3,500	1,000	13,125
9	15	20	–	–	–	–	–	–	–	–	35
10	10	4	–	–	–	–	–	–	–	–	14
11	60	100	50	1,800	1,000	–	500	700	–	100	4,310
12	1	–	–	–	–	–	–	–	–	–	1
13	50	15	150	–	–	1	–	–	–	–	216
15	2	1	–	1	–	–	–	–	–	–	4
20	–	–	–	1	1	–	–	–	–	–	2
23	–	–	3	–	–	1	–	–	–	–	4
24	–	–	–	–	–	–	–	1	–	–	1
25	–	–	–	–	4	1	–	–	–	–	5
28	–	–	–	2	–	1	–	–	–	–	3
29	–	–	–	–	1	–	–	–	–	–	1
30	–	–	–	–	–	1	–	–	–	–	1
32	–	–	–	–	–	–	–	1	–	–	1
36	1	–	–	–	–	–	–	–	–	–	1
40	–	–	1	–	–	–	–	–	–	–	1
44	–	1	–	–	–	–	–	–	–	–	1

(*) Days in which the commercial catch was observed.

TABLE A.4

Distribution of the sampling units, per species and individuals, trapped in fish-weir number 9, during ten days in June, 1980

Species Number	Sampling units (*)										Total
	17	18	19	20	21	22	23	24	25	26	
1	14	3	8	12	8	40	12	30	20	4	151
2	-	-	1	-	-	-	-	-	-	-	1
3	-	-	-	1	-	1	-	1	-	-	3
4	2	15	6	-	-	1	30	-	5	-	59
5	-	-	-	-	-	-	1	-	120	-	121
6	-	-	2	2	-	13	4	3	-	-	24
7	-	1	-	-	3	-	-	-	-	-	4
8	400	300	500	3,200	4,500	3,800	1,400	1,000	600	1,100	16,800
10	1	-	-	-	-	-	-	-	-	-	1
11	800	20	150	1,600	1,500	500	300	500	150	150	5,670
13	150	150	100	150	-	2	50	-	1	-	603
15	-	-	-	-	-	-	3	-	-	-	3
18	-	-	2	-	-	-	-	-	-	-	2
20	-	-	-	2	1	-	1	-	-	-	4
22	1	-	-	-	-	-	-	-	-	-	1
23	2	-	-	-	-	10	-	11	10	-	33
24	1	-	-	-	-	-	-	-	-	-	1
25	-	-	-	-	-	-	-	3	-	-	3
27	-	-	-	-	-	-	-	-	-	1	1
28	2	-	-	-	-	-	-	-	-	-	2
33	6	-	3	-	-	-	-	1	-	-	10
34	-	-	-	-	-	1	-	-	-	-	1
35	1	-	-	-	-	-	-	1	-	-	2
36	-	-	-	-	-	2	-	-	2	2	6
41	-	-	9	-	-	-	-	-	-	-	9
42	-	-	2	-	-	-	-	-	-	-	2

(*) Days in which the commercial catch was observed.

TABLE A.5

Distribution of the sampling units, per species and individuals, trapped in fish-weir number 10, during ten days in June, 1980

Species Number	Sampling units (*)										Total
	17	18	19	20	21	22	23	24	25	26	
1	15	10	13	15	11	–	3	–	10	9	86
2	2	2	–	–	–	–	–	–	–	–	4
3	–	2	–	4	–	–	9	80	2	6	103
4	15	2	–	10	–	–	6	4	15	8	60
5	2	1	–	–	–	–	6	4	3	–	16
6	–	–	4	–	–	–	–	–	1	–	5
7	10	–	–	2	6	–	10	–	–	–	28
8	2,000	1,000	1,000	3,900	3,200	3,600	400	300	1,000	800	17,200
11	500	300	900	950	800	400	100	50	300	400	4,700
13	100	30	150	120	4	–	20	–	1	–	425
15	–	–	–	–	–	–	–	2	1	2	5
16	–	–	1	–	–	–	–	–	–	–	1
20	–	–	–	1	1	1	–	–	–	–	3
22	10	–	–	–	–	–	–	–	3	–	13
23	8	–	–	18	16	–	–	8	15	10	75
24	–	2	–	1	–	–	–	–	–	1	4
25	–	–	–	–	16	–	9	–	–	–	25
26	–	–	–	–	23	–	–	–	–	–	23
28	–	–	–	4	–	–	–	–	1	–	5
29	–	–	–	1	–	–	–	–	–	1	2
32	–	1	–	–	–	–	–	–	–	–	1
33	2	–	–	–	–	–	–	–	5	1	8
35	–	1	–	–	–	–	–	–	–	–	1
36	1	1	10	–	–	–	1	–	–	1	14
37	–	–	–	–	–	–	–	–	–	2	2
39	–	–	–	–	–	–	–	–	–	1	1
40	–	–	–	–	–	–	1	–	–	–	1
41	–	–	–	–	–	–	5	–	–	–	5
43	–	–	–	–	–	–	1	–	–	–	1

(*) Days in which the commercial catch was observed.

TABLE A.6

Distribution of the sampling units, per species and individuals trapped in fish-weir number 16, during ten days in June, 1980

Species Number	Sampling units (*)										Total
	17	18	19	20	21	22	23	24	25	26	
1	28	27	19	17	13	18	55	35	22	47	281
2	-	-	1	-	-	-	-	-	-	-	1
4	5	-	8	12	-	6	-	6	-	12	49
5	-	-	-	1	-	-	-	-	-	1	2
6	-	1	-	-	-	-	-	-	-	-	1
8	1,400	1,800	2,000	3,000	1,600	400	150	200	500	1,200	12,250
11	500	2,200	1,500	900	1,000	900	400	50	-	700	8,150
13	100	2	-	-	2	10	13	4	1	-	132
14	-	1	-	-	-	-	-	1	-	-	2
15	-	-	-	1	-	-	-	-	-	-	1
18	3	-	-	-	-	1	-	-	-	-	4
19	-	-	-	-	-	-	-	-	-	1	1
20	1	-	-	-	-	-	-	-	-	-	1
23	-	7	8	10	-	-	15	8	20	22	90
24	-	-	-	-	-	-	1	-	-	-	1
25	8	9	-	3	-	8	12	3	3	1	47
35	-	-	5	-	-	1	19	-	-	-	25
36	-	2	-	2	-	5	2	-	1	-	12
38	-	-	-	-	-	-	-	-	-	2	2
40	-	-	-	-	1	1	-	-	-	-	2
43	-	-	-	-	-	-	1	-	-	-	1

(*) Days in which the commercial catch was observed.

APPENDIX B

Fitting a Truncated Lognormal Distribution

TABLE B.1

The fitting of a truncated lognormal distribution for the community of animals living in the area of fish-weirs at Almofala Beach (Ceará - Brazil).

i	ii	iii	iv	v	vi	vii	viii	ix		
r	$\log(r + \frac{1}{2}) = x^*$	f(r)	x*f(r)	$(x^* - \bar{x}^*)^2$ f(r)	$(x^* - \hat{u})/\sigma = z$	Φ(z)	$\hat{s}^*$Φ(z)	fe	goodness-of-fit	
									fe	f(r)
0	-0.30103	0	0	0	-0.7174	0.2358	13.58	(13.58)		
1	0.17609	8	1.40872	6.76443	-0.3129	0.3783	21.78	8.20	8.20	8
2	0.39794	4	1.59176	1.94709	-0.1248	0.4562	26.27	4.49	9.08	7
2 ⊣ 4	0.65321	3	1.95963	0.58721	0.0916	0.5359	30.86	4.59		
4 ⊣ 8	0.92942	3	2.78826	0.08288	0.3257	0.6255	36.01	5.15	5.15	3
8 ⊣ 16	1.21748	5	6.08740	0.07424	0.5699	0.7157	41.21	5.20	5.20	5
16 ⊣ 32	1.51188	5	7.55940	0.86632	0.8194	0.7939	45.71	4.50	4.50	5
32 ⊣ 64	1.80956	4	7.23824	2.03878	1.0718	0.8577	49.39	3.68		
64 ⊣ 128	2.10890	2	4.21780	2.05343	1.3255	0.9066	52.20	2.81		
128 ⊣ 256	2.40909	3	7.22727	5.17553	1.5800	0.9429	54.29	2.09	11.87	16
256 ⊣ 512	2.70969	3	8.12907	7.81557	1.8348	0.9644	55.65	1.36		
> 512	∞	4	-	-	-	1.0000	57.58	1.93		

APPENDIX C

Detailed Tables: Estimates of Jackknife Index

TABLE C.I

Pseudovalues of Jackknife estimates of the Shannon's index, for the community of animals living in the area of the fish-weirs at Almofala Beach (Acaraú-Ceará-Brazil).

Sampling units(k)	kth Sampling unit omitted	Pseudovalues $\hat{H}'_k{}^*$
1	0.806066	0.264999
2	0.809892	0.039230
3	0.803459	0.418808
4	0.804218	0.373993
5	0.799869	0.630569
6	0.796821	0.810410
7	0.802969	0.447693
8	0.798527	0.709778
9	0.796493	0.829773
10	0.796082	0.854019
11	0.805495	0.298676
12	0.795287	0.900925
13	0.800812	0.574936
14	0.791282	1.137238
15	0.800233	0.609100
16	0.796807	0.811279
17	0.805397	0.304443
18	0.798872	0.689392
19	0.798786	0.694519

Table C.1 (cont'd)

Sampling units (k)	kth Sampling unit omitted	Pesudovalues $\hat{H}_k'^*$
20	0.789933	1.216827
21	0.795801	0.870621
22	0.793571	1.002182
23	0.796362	0.837540
24	0.795119	0.910828
25	0.797064	0.796082
26	0.797313	0.781418
27	0.797812	0.751938
28	0.793870	0.751938
29	0.796157	0.849609
30	0.790297	1.195343
31	0.799829	0.632950
32	0.792147	1.086197
33	0.790428	1.187592
34	0.795669	0.878372
35	0.800050	0.619888
36	0.795914	0.863937
37	0.798195	0.729355
38	0.795229	0.904358
39	0.799319	0.663071
40	0.795959	0.861313
41	0.795534	0.886368

Table C.1 (Cont'd)

Sampling units (k)	kth Sampling unit omitted	Pseudovalues $\hat{H}'^{*}_{k}$
42	0.793395	1.012543
43	0.796502	0.829224
44	0.798987	0.682632
45	0.796376	0.836670
46	0.796823	0.810318
47	0.792609	1.058929
48	0.799226	0.668564
49	0.796415	0.834381
50	0.794367	0.955200
51	0.793931	0.980942
52	0.793822	0.987350
53	0.795242	0.903564
54	0.794312	0.958450
55	0.793158	1.026550
56	0.797045	0.797226
57	0.791280	1.137329
58	0.793731	0.992737
59	0.796589	0.824127
60	0.798569	0.707275

TABLE C.2

Pseudovalues of Jackknife estimates of the Shannon index, for fish-weir number 5, at Almofala Beach (Acaraú-Ceará-Brazil).

Sampling units(k)	kth sampling unit omitted	Pseudovalues $\hat{H}'^{*}_{k}$
1	0.776897	1.088591
2	0.772863	1.124896
3	0.794330	0.931693
4	0.881242	0.149488
5	0.817967	0.718956
6	0.829610	0.614171
7	0.823808	0.666386
8	0.791298	0.958979
9	0.810778	0.783660
10	0.804944	0.836164

TABLE C.3

Pseudovalues of Jackknife estimates of the Shannon index, for fish-weir number 6, at Almofala Beach (Acaraú-Ceará-Brazil).

Sampling units(k)	kth sampling unit omitted	Pseudovalues $\hat{H}'^{*}_{k}$
1	0.766541	0.815370
2	0.767269	0.808815
3	0.774211	0.746335
4	0.728844	1.154643
5	0.774231	0.746161
6	0.788777	0.615247
7	0.768566	0.797146
8	0.765569	0.824113
9	0.750611	0.958739
10	0.809183	0.431590

TABLE C.4

Pseudovalues of Jackknife estimates of the Shannon index, for fish-weir number 8, at Almofala Beach (Acaraú-Ceará-Brazil).

Sampling units(k)	kth sampling unit omitted	Pseudovalues $\hat{H}'^{*}_{k}$
1	0.668678	0.961486
2	0.680980	0.850769
3	0.671675	0.934515
4	0.644170	1.182064
5	0.694755	0.726796
6	0.691484	0.756236
7	0.716364	0.532311
8	0.698676	0.691509
9	0.779802	-0.038627
10	0.712698	0.565311

TABLE C.5

Pseudovalues of Jackknife estimates of the Shannon index, for fish-weir number 9, at Almofala Beach (Acaraú-Ceará-Brazil).

Sampling units(k)	kth sampling unit omitted	Pseudovalues $\hat{H}'^{*}_{k}$
1	0.744000	1.189503
2	0.769997	0.955526
3	0.771727	0.939960
4	0.782162	0.846045
5	0.844331	0.286521
6	0.839820	0.327117
7	0.789375	0.781126
8	0.782514	0.842876
9	0.759256	1.052200
10	0.804498	0.645022

TABLE C.6

Pseudevalues of Jackknife estimates of the Shannon indes, for fish-weir number 10, at Almofala Beach (Acaraú-Ceará-Brazil).

Sampling units(k)	kth sampling unit omitted	Pesudovalues $\hat{H}'^{*}_{k}$
1	0.738349	0.831789
2	0.746919	0.754659
3	0.699352	1.182765
4	0.765946	0.583422
5	0.766840	0.575375
6	0.815862	0.134171
7	0.735181	0.860303
8	0.728835	0.917418
9	0.742927	0.790593
10	0.738257	0.832623

TABLE C.7

Pseudovalues of Jackknife estimates of the Shannon index, for fish-weir number 16, at Almofala Beach (Acaraú-Ceará-Brazil).

Sampling units(k)	kth sampling unit omitted	Pseudovalues $\hat{H}'_k{}^*$
1	0.827564	0.988685
2	0.846313	0.819945
3	0.858246	0.712545
4	0.874927	0.562414
5	0.860570	0.691625
6	0.832246	0.946541
7	0.817148	1.082430
8	0.835059	0.921225
9	0.841196	0.865993
10	0.839202	0.883936