Acta Zoologica Academiae Scientiarum Hungaricae 58 (1), pp. 53–67, 2012

OBSERVATIONS ON SPECIES ABUNDANCE DISTRIBUTION IN FLY COLLECTIONS

IZSÁK, J.[1] and PAPP, L.[2]

[1]*Department of Systematic Zoology and Ecology, Eötvös Loránd University Pázmány Péter sétány 1/A, H-1117 Budapest, Hungary. E-mail: ijanos@elte.hu*
[2]*Beremend u. 43, H-1182 Budapest, Hungary. E-mail: flyer.papp@gmail.com*

Reports and methodological considerations relating species abundance distribution (SAD) in different multi-species communities are timely also today. We report detailed observations on SAD in a neutral fly community, sampled on cow pats in the period 2002–2005 in Hungary. The total sample consisted of 93 thousand individuals (faunistic observations were described in a former publication). We discuss among others observations on the change of frequency polygons with increasing sample size and advise further caution concerning the well-known shifting process, focussing on the discussion on 2-moving averaged abundance frequency polygons. We found a sample of about 20 thousand individuals to be yet insufficient and 30 thousand already sufficient to appear a modal abundance class. With larger sample size also a second internal mode was observable. The pool of the singleton and doubleton species was practically inexhaustible by increasing the sample size. A general conclusion is that a collection of the 93 thousand individuals may be yet insufficient to make a final conclusion on the SAD of the fly community sampled. As for observations on the shifting process working by increasing sample size, the shift of the abundance frequency polygon was significantly larger than expected. Remarkably, the shift of abundance class position of the species by increasing sample size in many cases differs significantly from the expected one. We report also other unexpected observations on the development of species abundance polygons by increasing sample size.

Key words: quantitative ecology, frequency polygon, mode, sample size, shifting process, singletons, Diptera

INTRODUCTION

An overall revived interest in the species abundance distribution (SAD) of different communities (GRAY *et al.* 2006, MCGILL *et al.* 2007, ULRICH *et al.* 2010 and numerous further papers available in electronic data bases). For finding theoretical background of SAD the book of HUBBELL (2001) gave a new impetus in the last decade. At the same time, a SAD model which is generally valid for almost all communities does not exist. This does not mean that some general traits of the SAD in different communities are missing. Such a trait is e.g. the inner mode of the frequency polygon if the sample size is sufficiently large. A similar trait is the large number (in a sense a frequency excess) of singleton and doubleton species even with very large sample sizes. After all, studying the SAD of formerly not studied communities remains actual. To attempt drawing a general conclusion seems to be timely in a later phase of the SAD investigations.

The aim of the present work is to describe some general properties of the species abundance distribution (SAD) in a neutral fly community sampled on pastures in Hungary. As for the neutrality, we suppose that, in the case of the present fly community relations between species can be neglected (cf. PAPP 1971). That is, this multi-species community is neutral and the species abundances should be taken for a set of independent realisations of a stochastic process leading e.g. to a lognormal SAD in the course of evolution (ENGEN & LANDE 1996). Neutrality does not play an essential role in our discussion.

Former data on SAD in fly collections are sketchy and sporadic. Among others this is our motivation to publish our findings. An early report originates from LAURENCE (1955), who found a poor fit of both the lognormal and logseries distributions to abundances in sample from the fly genus *Limosina* (see also WILLIAMS 1964: 44–45). There the flies were captured on cow pats during one year. The poor fit could be the consequence of the super-dominance of some species in the community. PETERSEN and MEIER (2003) reported some statistical traits of a large collection of Diptera. The authors do not discuss details on the SAD. Additionally, the collection sizes were tallied from the pinned collections reserved in museums. ARGEMÍ *et al.* (1999) reported a poor fit of log-series distribution to abundances in a collection of Drosophilidae. We found formerly a sufficient fit of truncated lognormal distribution and, in some cases, log-series distribution to Diptera collections (PAPP *et al*. 1997).

The new and considerably large collection of flies gave the possibility to analyse the SAD and numerous methodological issues. Besides this, the registration of species frequencies in rather small "elementary samples" and their appropriate combination made possible to generate practically arbitrary sized collections and follow systematically the effect of sample size on the observable SAD. As described by IZSÁK and PAPP (2011) in details, an elementary sample is a set of individuals captured on a single day and single pat or, in a few cases, two small pats. The elementary samples within a year were random arranged posteriorly.

One can gain information on the traits of the SAD among others by studying the frequency histogram or polygon of the species frequencies. As well-known by ecologists, the frequency histogram and polygon is shifting to the right as the sample size increases. Recalling the main elements of the shifting concept, let's consider the set of abundance classes the lower boundaries of which are e.g. $1, 2, 2^2, \dots$. Suppose that a given species abundance falls into the kth abundance class. When the sample size is doubled, then the new abundance will expectedly also be doubled and belongs to the $(k+1)$th abundance class. Consequently, by doubling the sample size the new frequency polygon can be obtained theoretically by shifting the polygon by one unit on a logarithmic scale to the right. Furthermore, in the first abundance class new species formerly hidden by the *veil line* will appear. According to the veil line principle and presuming a mode of the frequency polygon belonging to the whole community, by increasing the sample size a modal abundance class may appear first in the domain of the small frequencies. However, some doubts raise about the outlined shifting concept even if it corresponds to the experience in many cases (PIELOU 1969, 1975, PATIL & TAILLIE 1979, MAGURRAN 2004). Already PRESTON (1948) described a more realistic stochastic mechanism and DEWDNEY (1998) gave a more complete description. We decided to study the shifting process of the species abundances in more details.

To sum up, the main goal of this paper is to give information on the SAD of a large and important insect taxon such as flies. Besides this we report thought-provoking observations on the shifting process in connection with the development of the observed SAD.

MATERIALS AND METHODS

Study material

Flies were collected on cow pats in two pastures near Nagyiván, a village located in the Hortobágy National Park (HNP) in Hungary. Grazing on these pastures goes back decades. Such a long period was sufficient for all potential fly species to occupy this area and reach a relative stability of the community concerned. We collected flies between 2002 and 2005 and identified a total of 92 680 individuals. Number of individuals and registered species in the four years is given in Table 1. Taking into account our previous experience with diurnal and seasonal variations of population densities (ARGEMÍ *et al.* 1999, 2000), we collected the flies from the middle of July to the middle of August, and spent between 10 and 16 hours. More details see in PAPP (2007). The original data have been set up on the link of http://www.nhmus.hu/~lpapp. The number of registered species is 106. These data refer to a species-rich community. The fly community represents a single guild, because the species concerned are characteristic of the cow pat. (For types of relation to sites and food sources in the case of insects see e.g. NOVOTNÝ & BASSET 2000.) The species *Coproica lugubris* (HALIDAY, 1836) was super-dominant in all collections. The large number of species represented by very few individuals is also characteristic of all collections, see Table 1. The combination of collections gathered during some weeks compensates the fluctuations and succession changes in some days or weeks. On the other hand, a combination of one-year samples can reduce variations due to meteorological and similar yearly changes. For species accumulation curves and similarity patterns of collections and sub-collections of this fly community, see PAPP and IZSÁK (2011).

Statistical methods

We apply descriptive and mainly graphical methods based on abundance frequency polygons. Their use is not unusual, see e.g. figures in HUBBELL's monograph (2001). It should be mentioned that WILLIAMSON & GASTON (2005) described another related popular graphical method. According to our previous findings on Diptera (PAPP *et al.* 1997) our methodological starting point was that the observed SAD approximates the lognormal distribution. This is why our observations are based mostly on abundance frequency polygons of the log-transformed, namely the 2-based log-transformed species abundances.

We formed the abundance classes [1], [2–3], [4–7], … Similar ones were applied, for example, by KREBS (1989) and GRAY *et al.* (2006). The relatively narrow intervals [2^k, 2^{k+1}–1] are favourable for observing the distribution of small abundances; see also GRAY *et al.* (2006). The analysis of abundances falling in these classes is analogue to work with 2-based logarithms of the abundances. Sample variations were reduced by a well-known smoothing method, namely by halving the frequencies of abundances that fall onto the class boundaries, and classifying one half of the abundances to the lower and another one to the upper abundance class (PRESTON 1948, WILLIAMS 1964, GRAY *et al.* 2006). After determination of the abundance class frequencies we changed to 2-moving averages of the latters.

Table 1. Species abundance frequencies in the abundance classes. Possible modal frequencies in the 2004, 2005 and 2003–2004 collections are boldfaced.

Abundance class	2002	2003	2004	2005	2003–2004	2003–2005
1	13	22	12	16	18	10
2–3	8	12	17	6	13	19
4–7	4	12	6	7	11	10
8–15	3	8	**14**	**10**	**17**	11
16–31	3	7	8	8	9	3
32–63	4	5	5	7	8	17
64–127	1	3	5	6	5	7
128–255	0	4	5	9	6	9
256–511	0	3	2	5	6	2
512–1023	1	1	3	5	1	9
1 024–2 047	0	2	0	2	2	4
2 048–4 095	0	0	2	1	1	1
4 096–8 191	0	0	0	0	1	2
8 192–16 383	0	1	0	0	0	0
16 384–32 787	0	0	1	1	0	0
32 788–65 575	0	0	0	0	1	1
Individuals	1 146	21 009	29 525	41 000	50 534	91 534
Species	37	80	80	83	99	105

To follow up the development of the frequency polygon by increasing sample size gained by combining elementary samples, we formed gradually widening sub-collections with sizes growing approximately as the powers of 2. The sample sizes were 5 816, 11 043, 22 500, 46 319 (total collection consisted of 91 534 individuals, see Table 1). We were able to do this because species frequencies have been registered originally in the very small-sized elementary samples see in details in IZSÁK and PAPP (2011). From the latters, one can build up samples with almost arbitrary sample size. Moreover, we formed also "backward" series of sub-collections having sample size 5 317, 11 399, 22 799, 46 524 and 91 534. In addition, for further study of the variability of the polygon's shape, we partitioned the total collection into nearly equal-sized sub-collections A, B, C. The frequencies of the abundances in the abundance classes are given in Table 2.

RESULTS

Our observations on the SAD are based on the data in Tables 1–2 and averaged frequency polygons shown in Figs 1–4. The polygons in Figure 1a belong to the 2002 and 2003 collections with sizes 1 146 and 21 009, see Table 1. Both polygons have a monotone decreasing character. It is instructive that the shapes of the polygons are similar despite the enormous difference in the sample sizes. Furthermore, we can conclude that even the sample size of about 20 thousand individuals is yet insufficient to reveal a modal abundance class, visible in case of even larger sample size.

In case of polygons belonging to the 2004 and 2005 collections with sizes of 29 525 and 41 000 individuals (Fig. 1a) there appears a mode, namely at the fourth averaged abundance class. That is, in our case a sample size about 30 thousand

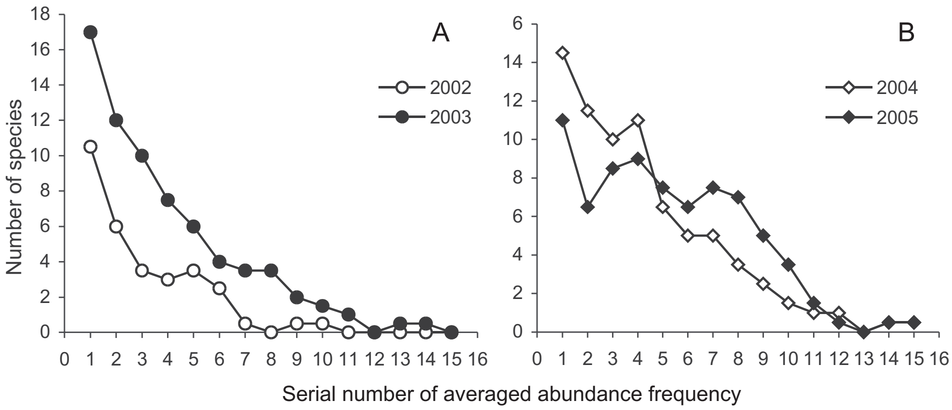

Fig. 1. Frequency polygons of 2-moving averaged abundance frequencies with the serial number of the average on the horizontal axis. For example, in the case of collection 2003 the first average is (22+12)/2 (see Table 1).

seems to be needed minimally to reveal a mode of the frequency polygon. The shape of the polygons is similar also in this case. With the 2005 polygon, belonging to the larger sample, there appears somewhat unexpectedly also a second mode. A further observation is the high number, in a sense the excess of singletons and doubletons, frequently reported also in case of other taxa (MAGURRAN & HENDERSON 2003).

The frequency polygons in Figure 2 belong to the combined collections 2003–2004 and 2003–2004–2005 with sample size 50 534 and 91 534. In contrast to the above cases, these collections are overlapping so far as the 2003–2004–2005 collection is an amplification of the 2003–2004 collection. The 2003–2004 polygon is shifted in the figure by two abundance classes to the right. This ensures a good parallelism with the 2003–2004–2005 polygon. However, this shift should correspond to an about *four-fold increase* in the sample size, while the real increase is only 91 534/50 534 = 1.8. This unexpected observation can not be interpreted on the basis of the outlined shifting process. We come back to this shift below. The excess of singletons and doubletons is observable also here in both cases.

As for data on the created sub-collections having each in turn a nearly doubled sample size (Table 2), we can observe the followings. In collections having sample size 5 816 and 11 043 we can observe a decreasing tendency of abundance class frequencies. Regarding samples with sample size 22 500 and 46 319 the series of these frequencies bears the marks of a mode about in frequency class [4–7] and [8–15] (Table 2). Similar can be said on the "backward" samples (Table 2).

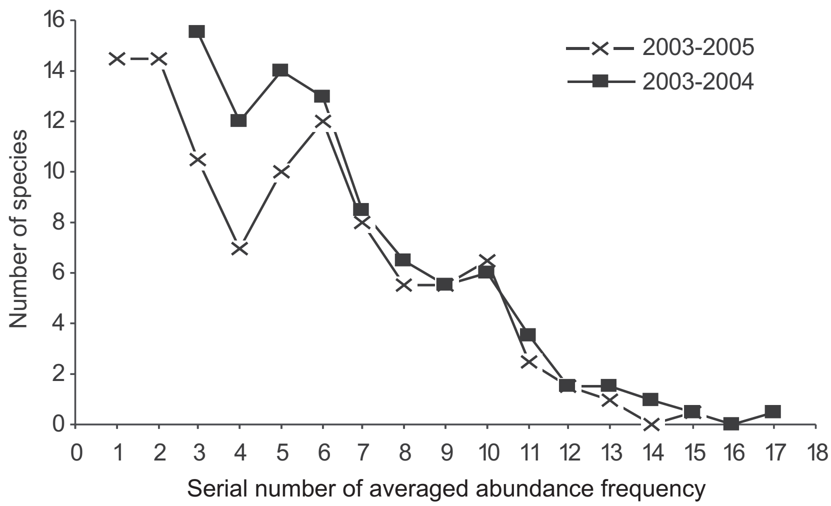

Fig. 2. Frequency polygons of 2-moving averaged frequencies relating to combined collections. The 2003–2004 polygon is shifted in the figure by two abundance classes to the right. Details see at Fig. 1.

Table 2. Basic data and species abundance frequencies in the abundance classes in the sub-collections. Data columns 1–5 and 6–10 stay for the forward and backward accumulated sub-collections, respectively. Some possible modal frequencies are boldfaced.

	individuals in the subsets												
											A	B	C
Abundance class	5 816	11 043	22 500	46 319	91 534	5 317	11 399	22 799	46 524	91 534	31 115	29 357	31 061
1	18	19	21	19	10	6	11	12	19	10	21	13	13
2–3	14	13	10	13	19	10	7	7	6	19	11	8	8
4–7	13	12	**14**	13	10	11	7	**9**	8	10	**12**	13	3
8–15	4	10	9	**14**	11	8	**10**	8	**9**	11	12	**14**	**12**
16–31	4	6	7	8	3	3	7	7	8	3	8	7	7
32–63	2	3	5	8	**17**	8	6	7	9	**17**	7	8	10
64–127	5	3	3	5	7	3	2	7	6	7	6	4	2
128–255	3	4	3	7	9	2	8	3	9	9	4	7	6
256–511	0	2	4	5	2	1	2	6	4	2	4	4	8
512–1023	1	1	2	2	9	0	1	2	6	9	2	1	3
1024–2047	0	0	0	1	4	0	0	2	1	4	1	1	2
2048–4095	1	0	1	1	1	1	0	0	2	1	1	2	0
4096–8191	0	1	0	1	2	0	1	0	0	2	0	0	0
8192–16383	0	0	1	0	0	0	0	1	0	0	0	0	0
16384–32787	0	0	0	1	0	0	0	0	1	0	1	1	1
32788–65575	0	0	0	0	1	0	0	0	0	1	0	0	0
Number of species	65	74	80	98	105	53	62	71	88	105	90	83	75

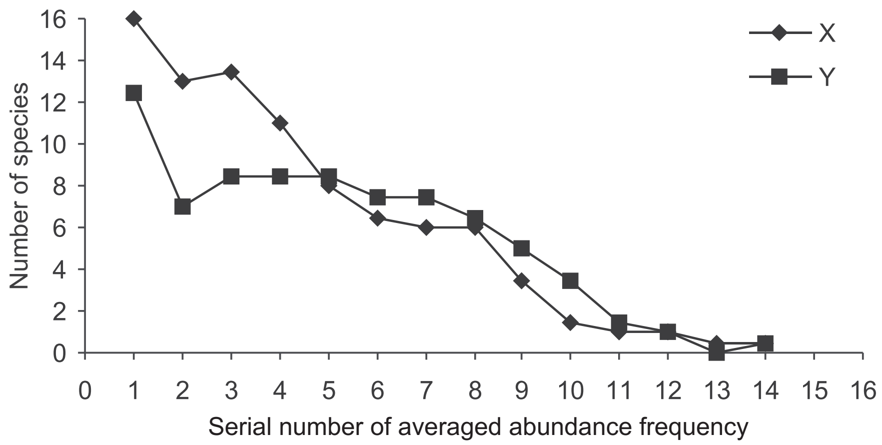

Fig. 3. Frequency polygons of 2-moving averaged frequencies. X and Y refer to the artificial collection containing the first 46 319 individuals and the last 46 524 individuals in the combined collection 2003–2004–2005. Details see at Fig. 1.

The smoothed frequency polygons belonging to the latter sample and the sample with of size 46 524 are shown in Figure 3. Both polygons resemble to those in Figure 1b, which belong to samples with 29 525 and 41 000 individuals, respectively.

According to the above data, a sample size of about 30 thousand seems to be a minimal value to manifest a mode of the frequency polygon. To confirm this supposition, we analysed also the three artificial collections A, B and C, each containing about 30 thousand individuals. The averaged frequency polygons are shown in Figure 4. One can realize that the polygons are more or less similar to those belong-

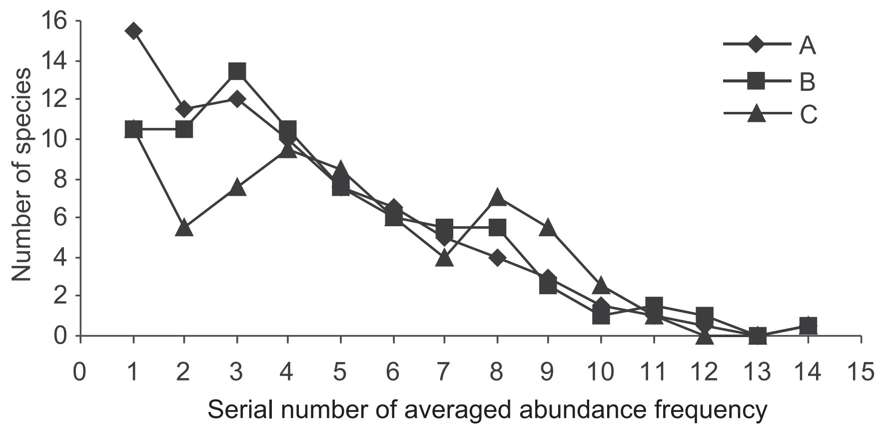

Fig. 4. Further frequency polygons of 2-moving averaged frequencies. A, B and C refer to artificial collections mentioned in the text. Details see at Fig. 1.

ing to collection 2004 and 2005 with sample sizes 29 525 and 41 000, respectively, see Figure 1. However, the signs of a mode can not be clearly observed.

Coming back to the shifting phenomenon, one can observe some further unexpected data. According to Table 1, in all of the three collections 2004, 2005 and 2003–2004 with sample size 29 525, 41 000 and 50 534 the interval [8–15] is a modal class. Taking into account the expected moving of species abundance by changing the sample size, the constancy of the modal class is surprising. Another unexpected observation is as follows. As mentioned above, the ratio of sample sizes in the 2003–2004–2005 and 2003–2004 collections is 1.812. Then, according to the shifting process concept, we should expect e.g. that most species belonging to abundance class [8–15] in sample with 50 534 individuals will occur in the class [16–31] in the sample with 91 534 individuals. At the same time, out of the 17 species concerned in the smaller sample only 2 ones fall into the class [16–31] in the almost double sized larger sample (see Table 3). From this table one can read out further similar unexpected cases. Concretely, forming the ratios of the observed/expected number of species abundances in the abundance classes concerned, we obtain the following data: [2–3]: 7/18, [4–7]: 4/13, [8–15]: 3/11, [16–31]: 2/17 (see above), [[32–63]: 6/9, [64–1273]: 1/8, [128–255]: 3/5, [256–511]: 0/6, [512–1023]: 3/6, [1024–2047]: 1/1, etc. We can establish that these unexpected transition data raise questions about the outlined shifting process. Table 3 informs in more details how species belonging to a given abundance class are scattered in abundance classes in another sample.

DISCUSSION AND CONCLUSIONS

To range our analyses and findings we have to take into account among others the notable book of HUBBELL (2001), in which the author discussed very extensively different factors and their role in the evolution of the SAD of various meta-communities. However, these factors prevail rather in long time periods not but in a single year or a couple of years. Anyway, detailed concrete analysis of changes in the SAD of different concrete taxa can not be replaced by theoretical considerations. Thus, the detailed study of changes in the SAD by increasing sample size is on the one hand important from a methodological point of view and it can serve for a basis of theoretical considerations, on the other. In addition, the study of the SAD in concrete taxa as much as possible may reveal new phenomena of the abundance distribution of species.

Numerous observations are published on the SAD in different taxa. A large collection of flies collected on pastures in Hungary provided a good opportunity to

Table 3. Transition data of species abundance classification. For example, the figure 7 indicates that out of the 18 singleton species in 2003–2004 collection only seven ones occur in the [2–3] abundance class in the 2003–2004–2005 collection.

	0	1	2-3	4-7	8-15	16-31	32-63	64-127	128-255	256-511	512-1023	1024-2047	2048-4095	4096-8191	8192-16363	16364-32787	32788-65575	Σ
0		2	3				1											6
1		8	7		2					1								18
2-3			9	4														13
4-7				6	3		1		1									11
8-15					6	2	3	2	4									17
16-31						1	6	2										9
32-63							6	1			1							8
64-127								2	3									5
128-255									1		5							6
256-511										1	3	2						6
512-1023												1						1
1024-2047												1	1					2
2048-4095														1				1
4096-8191														1				1
8192-16363																		0
16364-32787																		0
32788-65575																	1	1
Σ	0	10	19	10	11	3	17	7	9	2	9	4	1	2	0	0	1	

describe and ensure reference data on the SAD in a neutral fly community. According to our knowledge, the present description is the first comprehensive report on the SAD of fly communities. Results on different other taxa (catalogued recently by ULRICH *et al.* 2010) or theoretical considerations can not be substitute for a concrete new taxon such as flies. Further on, our methodological observations can contribute to general methodological issues. These observations were possible due to the possibility to construct nearly arbitrary sized sub-collections, as combinations of rather small elementary samples.

The description is based on the distribution of abundances in intervals with 2-power boundaries and 2-moving averaged frequency polygons of the logarithmed abundances, even if some authors criticized this procedure and suggest the application of rank abundance plots (NEKOLA *et al.* 2008). However, plotting the decreasingly ordered (log-transformed) abundances as function of their ranks is rather arbitrary and may form a somewhat deceptive impression on the original data set. We constructed abundance frequency polygons (see also e.g. in HUBBELL (2001): 37, etc.) instead of histograms. The application of polygons is advantageous in all cases when one wishes to plot simultaneously two or more frequency series.

The nearly arbitrary sized sub-collections make possible to follow up the development of the abundance frequency polygon by increasing sample size. The frequency polygon belonging to the total sample with 91 534 individuals characterizes most realistic the SAD of the original community. According to this polygon, beside a general decrease of frequencies there appears two modes. According to our knowledge, this type of SAD bimodality caused perhaps by the superposition of two unimodal SADs, is unknown in the literature. One can observe also with such a large sample size the excess of singletons and doubletons. This excess frequently occurs, among others in case of small-sized soil insects (OSLER & BEATTIE 1999). The biological background of this excess is unclear; see the instructive article of MCGILL (2003). At all, the question of rare species seems to be rather an ecological than a mathematical issue (GOLICHER *et al.* 2006). To eliminate the effect of this excess, MAGURRAN and HENDERSON (2003) attempted to separate the occasional species in a studied fish community, resulting in a nearly bell-shaped histogram of SAD for core species and a steeply decreasing histogram for the occasional ones (see Fig. 1 in the cited article). The separation of the community to occasional and core species may be realistic in some cases, even if NOVOTNÝ and BASSET (2000) reckon this method questionable. That is, the rare species should be reckoned as “tourists” (N.B.: this is not the case with our flies on cow pats) or satellite species in contrast to persistent or core species. The SAD of the two groups generally differ (ULRICH & ZALEWSKI 2006). Further studies would be required to analyze this issue in the fly data base.

On the other hand, as the number of species aside from the possibly small number of transient species is practically limited, with extraordinarily large collections the classes of small abundances would be empty and the left side of the abundance frequency polygon would be increasing. A reference point regarding this issue can be the estimated number of species missing from the sample. A simple estimate can be obtained from CHAO's first formula on the $\hat{S}_{tot}$ number of species in the total population (CHAO 1984): $\hat{S}_{miss} = \hat{S}_{tot} - \hat{S}_{obs} = f_1^2 / (2 f_2)$, where f_1 and f_2 are the number of singletons and doubletons in the sample. In case of collection 2003–2004–2005 $\hat{S}_{miss} = 10 \times 10 / (2 \times 14) = 4$. (Estimates for the other collections are: $\hat{S}_{miss}$ (2003) = 30, $\hat{S}_{miss}$ (2004) = 12, $\hat{S}_{miss}$ (2005) = 32, $\hat{S}_{miss}$ (2003–2004) = 18). The result suggests that the missing species can not influence significantly the SAD in the largest sample. One should note that these richness estimators serve only for a first reference. That is, a further change of the frequency polygon by increasing the sample size can not be expected. Naturally, the above estimation may be insufficient and the exhaustion of the species set is perhaps not a practical possibility. In addition, there is not a clear relationship between the proportion of singletons and completeness of the sample (PETERSEN & MEIER 2003).

As for the development of the polygon's shape by increasing sample size up to about 20 thousand individuals we obtained monotone decreasing frequency polygons (Fig. 1a). A modal abundance class appears approximately at 30 thousand or more individuals in the sample. It is worth here mentioning ULRICH's (2001) observations, who found by means of numerical experimentation that in case of a community of 100 species about 50 000 individuals is necessary for the identification of an abundance distribution. According to our observations data on SAD in fly collections even with a sample size of 46 thousand should be taken by caution and frequency polygons with such collection sizes vary significantly (see Table 2, Figs 3–4). On the other hand, taking into account the variable conditions with different taxa, our observations do not differ essentially from ULRICH's data. It should be noted that in case of a fish collection consisting of cca. 96 thousand individuals MAGURRAN and HENDERSON (2003) obtained a very similar result on the SAD, see figure 1a in their paper. With the mentioned sample size also an expressed second internal mode appears. Possibly we have to suppose the existence of two subpopulations, in which case a lognormal distribution and single mode of the combined population can be excluded (STORCH & ŠIZLING 2008).

We do not discuss the fit of different distributions to the frequency series. The lognormality of the SAD can be excluded in the present case. In case of the smaller samples where the abundance frequency polygon is decreasing, fitting among others Zipf – Mandelbrot distribution (IZSÁK 2006), (truncated) lognormal distribution, Poisson lognormal distribution (IZSÁK 2008) is often successful.

However, even if the results seem to be apparently sufficient, it is possible that the good fit should be considered an artefact caused by missing a significant part of the original distribution.

It seems realistic to interpret the appearance of the mode and generally, the change of the polygon's shape by increasing sample size by the shift of the frequency polygon. However, the observed unexpected traits of the shift do not correspond to the general concept of shifting. For example, a shift by two units instead of one unit was needed to achieve the best parallelism between the polygons belonging to collections 2003–2004 and the amplified collection 2003–2004–2005. This shift corresponds to a four-fold increase in the sample size. At the same time, the real increase was only 1.8. Observing the shifting process on the species level we found further curious phenomena. The features of the discussed transition table (Table 3) demonstrating convincingly the fuzzy nature of the shifting process. This confirms our opinion that the shifting concept is not applicable in the recent form to forecast the changes in the SAD by increasing sample size.

Some further unexpected phenomena such as the constancy of the mode by increasing sample size and the appearance of a second mode are discussed in the Results section.

In general, we think that separate discussion of the SAD in case of species-rich and species-poor community seems to be reasonable (ALONSO & MCKANE 2004). Another possibility is that a collection of about 92 thousand individuals is yet insufficient to establish a final conclusion on the SAD in the present fly community. One should note that taking into account the costs of collecting it is a considerable undertaking to produce such a large collection.

*

Acknowledgements – The study was supported by the Hungarian Scientific Research Fund (OTKA), No. T 043168 for J. I. and L. P. We thank three anonymous reviewers for their comments and advice, and Mrs ANIKÓ KOROKNAI for her assistance in informatics.

REFERENCES

ALONSO, D. & MCKANE, A. J. (2004) Sampling Hubbell's neutral theory of biodiversity. *Ecology Letters* **7**: 901–910.

ARGEMÍ, M., MONCLÚS, M., MESTRES, F. & SERRA, L. (1999) Comparative analysis of a community of drosophilids (Drosophilidae; Diptera) sampled in two periods widely separated in time. *Journal of Zoological Systematics and Evolutionary Research* **37**: 203–210.

ARGEMÍ, M., BÄCHLI, G., MESTRES, F. & SERRA, L. (2000) Analysis of the diurnal activity patterns of six species of *Drosophila* (Diptera, Drosophilidae) at Bordils (North East of Spain). *Mitteilungen der schweizerischen Entomologischen Gesellschaft* **73**: 337–348.

CHAO, A (1984) Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics* **11**: 265–270.

DEWDNEY, A. K. (1998) A general theory of the sampling process with applications to the "veil line". *Theoretical Population Biology* **54**: 294–302.

ENGEN, S. & LANDE, R. (1996) Population dynamic models generating the lognormal species abundance distribution. *Mathematical Biosciences* **132**: 169–183.

GOLICHER, D. J., O'HARA, R. B., RUÍZ-MONTOYA, L. & CAYUELA, L. (2006) Lifting a veil on diversity: a Bayesian approach to fitting relative-abundance models. *Ecological Applications* **16**: 202–212.

GRAY, J. S., BJØRGESAETER, A. & UGLAND, K. I. (2006) On plotting species abundance distributions. *Journal of Animal Ecology* **75**: 752–756.

HUBBELL, S. P. (2001) The Unified Neutral Theory of Biodiversity and Biogeography. *Monographs in Population Biology* **32**. Princeton University Press, Princeton, New Jersey

IZSÁK, J. (2006) Some practical aspects of fitting and testing the Zipf-Mandelbrot model. A short assay. *Scientometrics* **67**: 107–120.

IZSÁK, J. & PAPP, L. (2011) Species accumulation curves and similarity traits of a species-rich fly (Diptera) community. *Acta Zoologica Academiae Scientiarum Hungaricae* **57**(1): 85–94.

IZSÁK, R. (2008) Maximum likelihood fitting of the Poisson lognormal distribution. *Environmental and Ecological Statistics* **15**: 143–156.

KREBS, C. J. (1989) *Ecological methodology*. Harper Collins Publishers, 630 pp.

LAURENCE, B. R. (1955) The ecology of some British Sphaeroceridae (Borboridae, Diptera). *Journal of Animal Ecology* **24**: 187–199.

MAGURRAN, A. E. (2004) *Measuring biological diversity.* Blackwell Publishing Company, 268 pp.

MAGURRAN, A. E. & HENDERSON, P. A. (2003) Explaining the excess of rare species in natural species abundance distributions. *Nature* **422**: 714–716.

MCGILL, B. J. (2003) Does Mother Nature really prefer rare species or are log-left-skewed SADs a sampling artefact? *Ecology Letters* **6**: 766–773.

MCGILL, B. J., ETIENNE, R .S., GRAY, J. S., ALONSO, D., ANDERSON, M. J., BENECHA, H. K., DORNELAS, M., ENQUIST, B. J., GREEN, J. L., HE, F., HURLBERT, A. H., MAGURRAN, A. E., MARQUET, P. A., MAURER, B. A., OSTLING, A., SOYKAN, C. U., UGLAND, K. I. & WHITE, E. P. (2007) Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecology Letters* **10**: 995–1015.

NEKOLA, C. J., ŠIZLING, A. L., BOYER, A. G. & STORCH, D. (2008) Artifactions in the log-transformation of species abundance distributions. *Folia Geobotanica* **43**: 259–268.

NOVOTNÝ, V. & BASSET, Y. (2000) Rare species in communities of tropical insect herbivores: pondering the mystery of singletons. *Oikos* **89**: 564–572.

OSLER, G. H. R. & BEATTIE, A. J. (1999) Relationships between body length, number of species and species abundance in soil mites and beetles. *Pedobiologia* **43**: 401–412.

PAPP, L. (1971) Ecological and production biological data on the significance of flies breeding in cattle droppings. *Acta Zoologica Academiae Scientiarum Hungaricae* **17**: 91–105.

PAPP, L. (2007) A study of the cow pat Diptera on the Hortobágy, Hungary. *Folia entomologica hungarica* **68**: 123–135.

PAPP, L., IZSÁK, J. & ÁDÁM, L. (1997) Dipterous assemblages of sheep-run droppings: number of species observed, estimated and generated by simulation. *Acta Zoologica Academiae Scientiarum Hungaricae* **43**: 191–205.

PATIL, G. P. & TAILLIE, C. (1979) Species abundance models, ecological diversity, and the canonical hypothesis. *Bulletin of the International Statistical Institute* **44**: 1–23.

PETERSEN, F. T. & MEIER, R. (2003) Testing species-richness estimation methods on single-sample collection data using the Danish Diptera. *Biodiversity and Conservation* **12**: 667–686.

PIELOU, E. C. (1969) *An introduction to mathematical ecology.* Wiley, New York, 286 pp.

PIELOU, E. C. (1975) *Ecological diversity.* Wiley, New York, 176 pp.

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

STORCH, D. & ŠIZLING, A. (2008) The concept of taxon invariance in ecology: do diversity patterns vary with changes in taxonomic resolution? *Folia Geobotanica* **43**: 329–344.

ULRICH, W. (2001) Models of relative abundance distributions I: model fitting by stochastic models. *Polish Journal of Ecology* **49**: 145–157.

ULRICH, W. & ZALEWSKI, M. (2006) Abundance and co-occurrence patterns of core and satellite species of ground beetles on small lake islands. *Oikos* **114**: 338–348.

ULRICH, W, OLLIK, M. & UGLAND, K. I. (2010) A meta-analysis of species-abundance distributions. *Oikos* **119**: 1149–1155.

WILLIAMS, C. B. (1964) *Patterns in the balance of nature*. Academic Press, London, 332 pp.

WILLIAMSON, M. & GASTON, K. J. (2005) The lognormal distribution is not an appropriate null hypothesis for the species-abundance distribution. *Journal of Animal Ecology* **74**: 409–422.

Revised version received April 27, 2011, accepted September 22, 2011, published May 10, 2012