ACTA SOCIETATIS
BOTANICORUM POLONIAE
Vol. 53, nr 1: 87-98
1984

Species frequency distribution as a function of the surface area

ANNA JUSTYNA KWIATKOWSKA

Department of Phytosociology and Plant Ecology, Institute of Botany, Warsaw University, Al. Ujazdowskie 4, 00-478 Warszawa, Poland

(Received: December 21, 1982. Revision accepted: June 7, 1983)

Abstract

Paper deals with the results of statistical analysis of the type of frequency distribution of species occuring in the field layers of two forest phytocoenoses. In the both cases frequency distributions were ranged out for the surface area of 1, 2, 4, 8, 16, 32 and 64 m^2. The types of frequency distributions were determined on the grounds of the values of Fisher's and Pearson's K coefficients. Analysed distributions were classified into Pearson's system. Also the size of the sample plot at which the empirical frequency distributions were symetrical, from the statistical point of view, nad where they were U-shaped was determined.

Key words: homogeneity, frequency distribution, minimum area

INTRODUCTION

In geobotany the frequency index has been often used as a measure of probability of finding the species on the investigated area. In the contemporary investigations the term of frequency is referred to the organization level of ecosystem. The scale size is therefore determined by the phytocoenose biochore, thus calculated frequency value defines the probability of finding the species in the patch. This probability (expressed in per cent) is estimated on the grounds of the series of surface samples of the definite size. The histogram of frequency distribution is constructed by calculating frequency of all the species, occuring in the patch and later grouping species in frequency classes. Number of species belonging to the particular class is expressed as the percentage of the total species number in the sample. The highest histogram peak is formed mostly by the numerous group of species of the first frequency class, rare in the patch. The high participation of the species common on the investigated area results in the second, lower histogram peak. As the histogram peaks occur in the first and the last frequency class, the histogram itself is U-shaped. The species are usually grouped according to Raunkiaer (1918) in five frequency classes, each including 20 per cent interval.

Conformity of the frequency distribution of species occuring on the investigated area and Raunkiaer's normal frequency distribution of the U type (Raunkiaer 1918) is one of the earliest used criteria of floristic homogeneity of the area (Dagnelie 1962, Daget and Godron 1982). It happens that the criterion of the type of the frequency distribution is used not only in the spatial meaning, concerning the individual objects (phytocoenoses), but also in the abstract one based on the probability of the species occurrence in the type (phytocoenon). For example Guinochet (1954, 1955, 1957), Géhu and Géhu (1969, 1975a and b) Géhu (1976) and De Foucalt (1979) have used the method of analysis of the histogram of species occurrence, as the criterion of floristic homogeneity in the phytosociological tables.

Dagnelie disagreed with that (1962). He used Raunkiaer's method of the analysis of the type of frequency distribution on the grounds of only one phytocoenosis, but holding investigations on the different-sized sample plots. The author did not discuss the fact, that the general populations for which the probabilities of finding the element were estimated, considerably differ from each other. Thus he is wrong thinking that his results disqualify the fitness of Guinochet's criterion (1954).

The results of Dagnelie's investigations, as well as those obtained for the same object by the others (Calléja et al. 1962) are convergent. Value of the sample plot size, for which the frequency distribution is symmetrical from the statistical point of view, corresponds with the size of the representative area calculated by other methods.

The convergence of the sizes of plots, which according to the well known criteria are homogenous, observed by the author, using various methods based on the different structural indices of the phytocoenose inspired this paper. In case when the results are not casual this fact is very interesting from the methodical as well as cognitive point of view. The investigations of Calléja et al. (1962) are pioneer and were held on the grounds of one particular type of the phytocoenose (sod), that is why it seems to be justified to start analysis which will answer the question, wheather the results obtained become true in the cases of other phytocoenoses.

The aim of the paper is to determine wheather in the analysed cases (two forest phytocoenoses) following relationships, proved to be true by French authors, are confirmed:

1) type of the frequency distribution — plot size,

2) values of statistical characteristics of the empirical frequency distributino — plot size,
3) degree of homogeneity (on the grounds of Raunkiaer's criterion) — distribution — plot size,

Besides the author was to determine wheather the value of the plot size, when the distribution is symmetrical from the statistical point of view, is convergent with the size of the representative area calculated by other methods. The results of the work for Master degree of Dołgopoł (1971) were used.

MATERIAL AND METHODS

Two different phytocoenoses in the Białowieża Primeval Forest: oak-hornbeam forest *(Tilio-Carpinetum)* and oak forest *(Potentillo albae-Quercetum)* were the objects of the investigation. The oak-hornbeam forest investigation plot was located in Zwierzyniec forest district on the territory of the partial reserve, in 452 division. In the stand tree cover was about 80 per cent, *Carpinus betulus* dominated. Shrubs were weak and small (10 per cent cover). The densely of the field layer was also about 80 per cent and there occured 89 species of the vascular plants. The second (oak forest) plot was placed in Hajnówka forest district, 442 division. It was described in details by Matuszkiewicz and Wydrzycka (1972).

The base of the analysis were the lists of all the species occuring in the sample plots. The sample plots were all the squares, their sides of 1 m, 512 in each phytocoenose. They were placed in the systematic sample design according to Greig-Smith's lattice (1952). The work is a comparative one, so regarding the collection and elaboration of the samples all Dagnelie's methods have been used. The plot sizes (1, 2, 4, 8, 16, 32, 64 m^2), as well as the way of doubling and grouping the samples were accepted (Matuszkiewicz and Wydrzycka 1972). This type of the design enabled to group the samples in the lattices with the increasing size of the elementary plot. It also allowed the quick analysis of the frequency distribution types due to the size of the plot. To obtain this information the values of all the frequencies of the species were calculated. Then the species were gouped in eight frequency classes (according to Dagnelie (1962) the interval width is 12.5 per cent). Then the frequency table and exactly corresponding with it the frequency histogram were constructed.

Further the frequncy tables were regarded as the statistical series of the attribute — probability of the occurrence of the species of definite frequency on the sample area of the different but definite size. The basic statistic were calculated for the series and they gave the full distribution characteristics. To obtain these parameters it was necessary to calculate the distribution moments of the first, second, third and the fourth rank, and later on the grounds of these — the arithmetical mean ($\bar{x}$), variance (s^2), β_1, β_2, γ_1 coefficients (Yule and Kendall 1966) and K Pearson's coefficient (1902).

β_1 coefficient is a measure of the distribution disymmetry and amounts to O for the symmetrical distribution. Fisher's γ_1 coefficient = $= \sqrt{\beta_1}$ has similar meaning. β_2 coefficient is mostly used to calculate kurtosis of the one-peak distribution. The standard values of β_1 as well as β_2 equal 3. The curves with values smaller than 3 are called platykurtic (wide) curves and those with values bigger than 3 — leptokurtic (narrow) curves. The degree of the distribution deviate from the symmetry is well expressed by the K coefficient usually used in the following formula:

$$K = \frac{\beta_1 (\beta_2 + 3)^2}{4 (4 \beta_2 - 3 \beta_1)(2 \beta_2 - 3 \beta_1 - 6)} .$$

The use of this coefficient enables to classify a particular empirical distribution in to Pearson's system of corresponding curve types.

The most important fact is that for the symmetrical distribution K coefficient equals O. Thus the distribution curve belongs to the type II or VII in Pearson's system. For distributions of the type II β_2 coefficient is smaller than 3. When β_2 ranges between 3 and 1.8 — the distribution curve is bell-shaped. And when β_2 is smaller than 1.8 the distribution curve is of the U type.

On the grounds of the values of coefficients β_1, β_2 and K it is possible to determine precisely wheather the distribution is symmetrical from the statistical point of view (β_1 and K=O) or wheather it is U-shaped ($\beta_2 < 1.8$).

The analysis of the type of the relations between the investigated characteristics and the area size was carried out by the least squares method and approximation of the curve type on the grounds of empirical data (Yule and Kendall 1966).

RESULTS

Figure 1 (for the oak forest) and Fig. 2 (for the oak-hornbeam forest) illustrate the results of the statistical analysis of the relation empirical distribution of the frequency — plot size. In both cases frequency histograms for the species occurring on the surface area of 1 m^2 are of the L type (Figs. 1, 2). On the sample plots of this size there is predominance of the rare species of the lowest frequency class over the common species, those for which probability of finding them in the patch exceeds 0.875. The histogram of the L type begin to change into histograms of the U type when there are 8 and 16 m^2 sample plots. It is caused by the increase in percentage of the common species in the sample. According to Raunkiaer's criterion (1918) these histograms correspond with the normal frequency distribution of the U type.

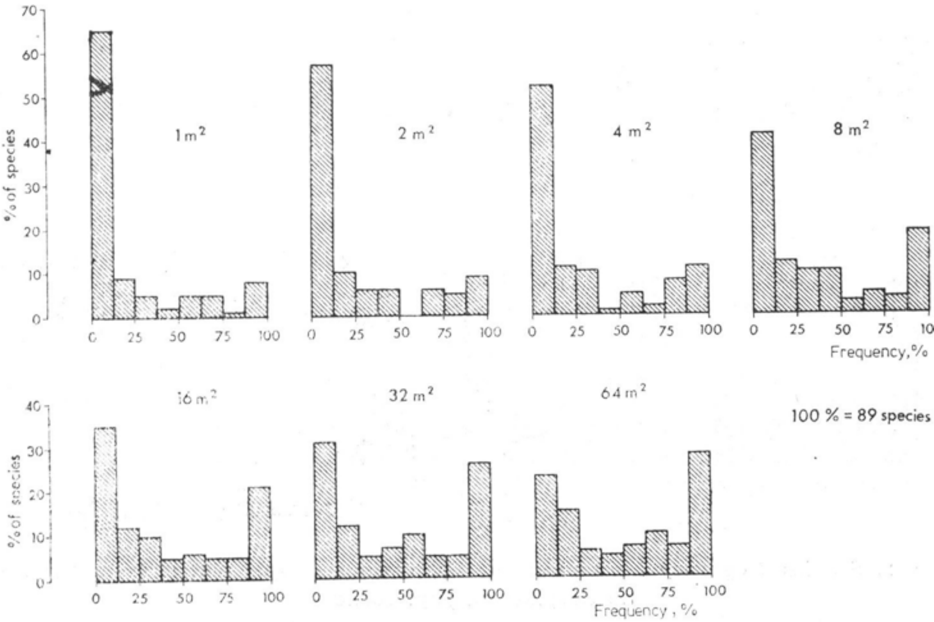

Fig. 1. Species frequency histograms for the various sample plot sizes in the *Potentillo albae-Quercetum* phytocoenose

Equalization of the species number in the first and the last frequency classes occurs in the oak forest for the 32 m^2 plot (Fig. 1), and in the oak-hornbeam forest — for the 64 m^2 plot (Fig. 2). Histograms are clearly two-peaked, symmetrical, U-shaped. Assuming the criterion of homogenity is true it can be stated that these plots (32 and 64 m^2) fullfy the condition of the floristic homogeneity of the analysed patches.

This relation was elaborated with the help of statistical methods and was expressed quantitatively as a function of the surface area for such distribution parameters as $\bar{x}$, s^2, β_1, β_2, γ_1 and K. Figures 3-8 illustrate the values of the statistics for the different sizes of sample plots. The curves were ranged out in a half-logarithmic scale. The values of plot size (p) as the independet variable were transformed according to a function $x = \log_2 p$. The dependent variable values were expressed on the line scale.

Figure 3 illustrates the average species frequency as a function of the surface area (plot size), Fig. 4 — values of γ_1 coefficient. In both phyto-

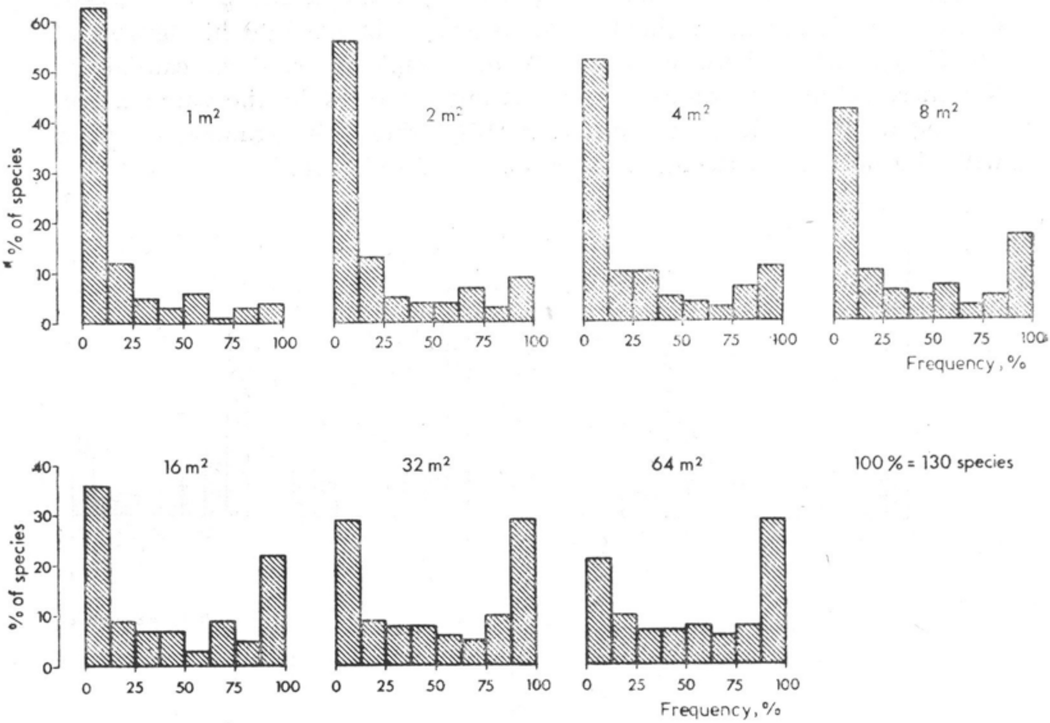

Fig. 2. Species freguency histograms for the various sample plot sizes in the *Tilio-Carpinetum* phytocoenose

coenoses observed values deviate non-significantly from the straigth line approximated on the grounds of empirical data by the least squares method. Function relation can be expressed in the p value interval from 1 to 64 by the linear equatation $y = \pm ax + b$. All that results in the dependence of the values of the mean species frequency in the phytocoenose as well as the values of γ_1 coefficient on the plot size and is proportional to its logarithm. While comparing two phytocoenoses it can be stated that in the analysed range of plot sizes mean frequency values of the species in the oak forest are, as a rule, higher and increase more quickly than in the oak-hornbeam forest.

Graphs shown on the Fig. 5 illustrate relation between frequency variance and plot size. In both phytocoenoses observed values deviate non-significantly from the parabole line, approximated by the least square method. Function relation in the analysed interval of p values is expressed by the equatation $y = -ax^2 + bx + c$. That indicates that variance of the species frequency in the phytocoenose is negatively proportional to the square of the logarithm of the elementary plot surface. The same type of relation concerns K coefficient (Fig. 6).

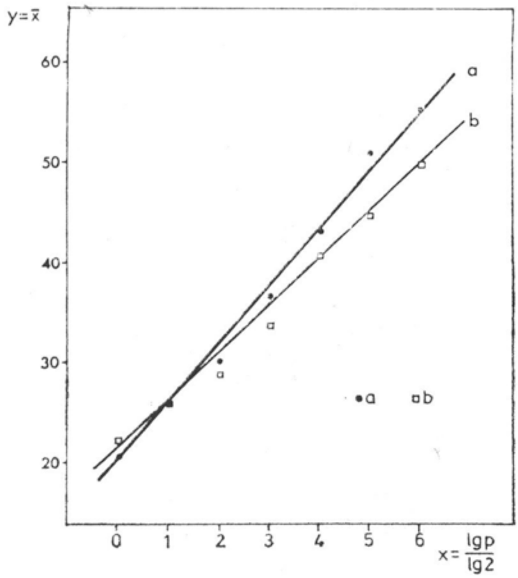

Fig. 3. Mean frequency ($\bar{x}$) of the field layer species — sample plot size (p), curves: a — *Potentillo albae-Quercetum*, b — *Tilio-Carpinetum*

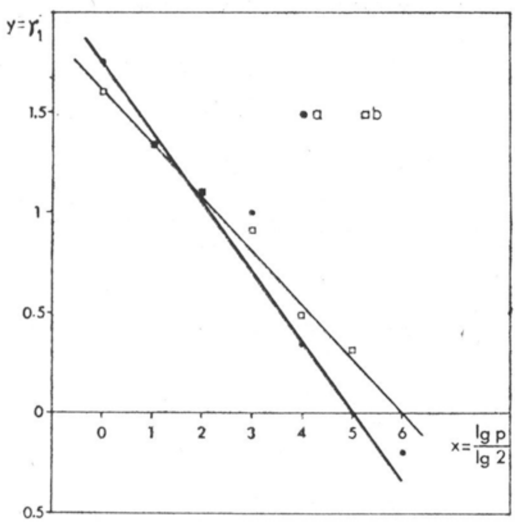

Fig. 4. γ_1 coefficient value of the frequency distribution of the field layer species — sample plot size (p), curves a and b as above

In both phytocoenoses also β_1 and β_2 coefficient values devite non--significantly from the parabole line (Figs. 7 and 8). Function relations can be thus expressed by the equatation $y = ax^2 + bx + c$. Values of the measures of asymmetry and flattening of the curve are therefore proportional to the logarithm of the elementary plot size.

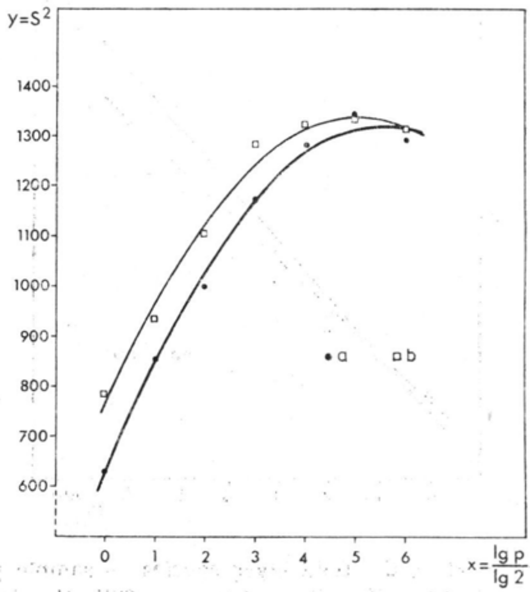

Fig. 5. Frequency variance (s^2) of the field layer species — sample plot size (p), curves a and b as above

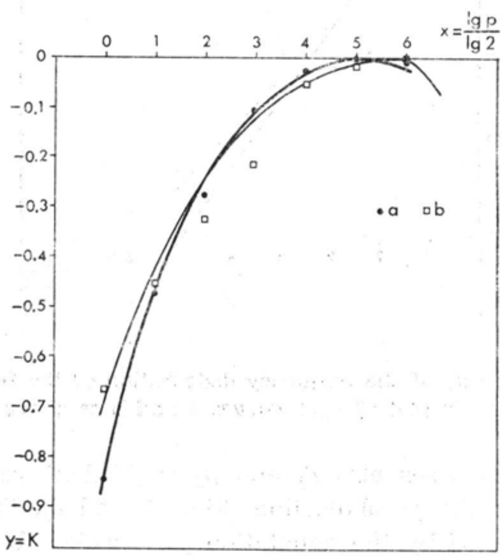

Fig. 6. K Pearson's coefficient value — sample plot (p) size, curves a and b as above

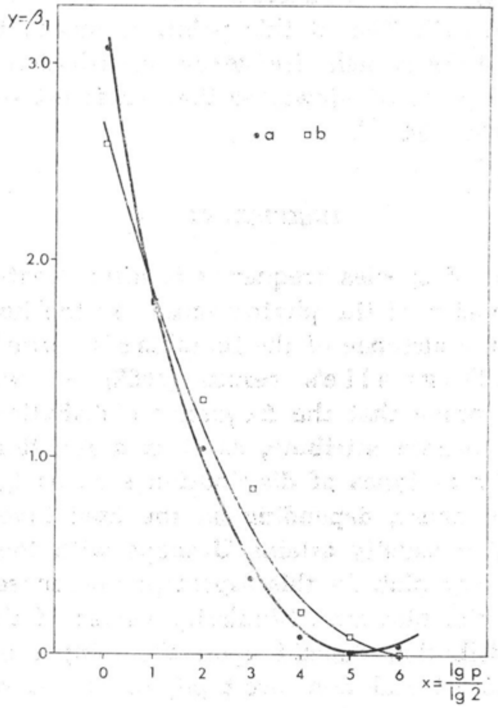

Fig. 7. β_1 coefficient value of the frequency distribution of the field layer species — sample plot size (p), curves a and b as above

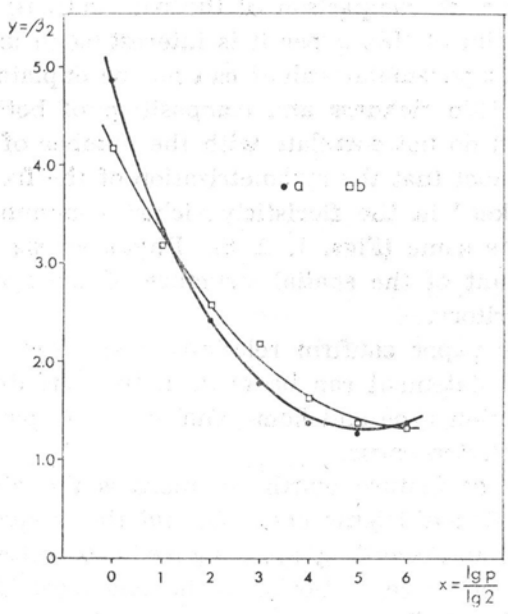

Fig. 8. β_2 coefficient value of the frequency distribution of the field layer species — sample plot size (p), curves a and b as above

For β_1 and K coefficients function minimum is O, it indicates the symmetry of the distribution at this point. It means that there exists the size of the plot for which frequency distribution is symmetrical from the statistical point of view; for the oak-forest — 32 m^2, for the oak-hornbeam forest — 64 m^2.

DISCUSSION

The distribution of species frequency is often treated as a criterion of floristic homogeneity of the phytocoenose. As the homogeneity minimum one considers occurrence of the Raunkiaer's normal U-typed distribution. However, Dagnelie's results (1962) — which were also confirmed here — prove that the frequency distribution type is not an independent phytocoenose attribute, as it is a function of the sample plot size. The different types of distribution such as L, U, J can occur in the same phytocoenose, depending on the used investigation surface size. The distribution usually attains U shape with the increase in the size of the elementary plot. In this aspect phytocoenose homogeneity is also a function of the plot size. Similarily values of the parameters of the frequency distribution depend upon the sample plot size. For all the investigated, forest and non-forest phytocoenoses relations seem to be the same. As the sample plot size increases, values of the coefficients of variance and asymmetry change according to the parabole equatation. The difference between the phytocoenoses lies only the analyzed parameter values. Although comparison of the oak and oak-hornbeam forest has not been the aim of this paper it is interesting to underline that the reasons of different parameter values can not be explained by the differences in the floristic richness and composition of both phytocoenoses. The stated relation do not correlate with the number of species in these communities. The fact that the symmetrization of the frequency distribution occurs "quicker" in the floristicly richer community of the oak forest indicates the same (Figs. 1, 2, 6). Phytocoenose homogeneity is therefore a resultant of the spatial structure of the species populations occupying this territory.

Results of this paper confirm relations discovered by the French authors. Thus the statement can be formed, that the dependence of the frequency distribution type and homogeneity on the plot size is a structural rule of the phytocoenose.

The phytocoenose feature worth comment is the plot size at which values of β_1 and K coefficients equal O, and the frequency distribution is realy symmetrical. According to Dagnelie's criterion (1962) phytocoenose at this plot size is floristicly homogenous. The relation type between the analysed distribution parameters (linear or parabolic) and the logarithm of the plot size is not so important on the merits. Its determination is to some degree faulty, the error comes from the systematic sample collection design. However it is worth to show that the results described above agree with those of Dagnelie (1962).

The plot size which according to the accepted above criterion is homogenous and the size of the representative area are similar. For example the representative area size in the oak forest calculated by the differential quotient method amounted to 27.6 m^2 (Matuszkiewicz and Wydrzycka 1972); the plot size were the random type of spatial distribution of the species prevailed was 32 m^2 (Kwiatkowska 1972). But the plot size at which phytocoenose is regarded to be homogenous can be considered as the optimal sample plot size — representative floristicly for the individual phytocoenose. The analysed phytocoenoses differ mainly in the plot size at which the distribution is statisticly symmetrical. Obviously, frequency distributions in the oak and oak-hornbeam forests differ in calculated values of the parameters and their function direction components that define the parameters dependence upon the plot size.

The use of this convergence in the determination of the representative area should be based mostly on the calculations of β_1 coefficient, as calculating K coefficient is very laborious. If it is impossible to obtain $\beta_1 = O$ for any of the plot sizes it can be done indirectly by the analysis of γ_1 coefficient graph. γ_1 coefficient is lineary inversely proportional to the logarithm of the plot size. It is very easy to find on the graph the plot size where regression line intersects x axe.

The quickest (but not very accurate) method of determination of the frequency distribution symmetry is to find the point where both the first and the last frequency classes do not differ in number. That estimation can be even precise, if there is no peak in the middle frequency classes.

REFERENCES

Calléja M., Dagnelie P., Gounot M., 1962. Etude statistique d'une pelouse à *Brachypodium ramosum*. Bull. Serv. Carte Phytogéog. Ser. B 7: 44-111.

Daget Ph., Godron M., 1982. Analyse frequentielle d'écologie des espèces dans les communautés. Masson, Paris.

Dagnelie P., 1962. La distribution des frequences des espéces. Bull. Serv. Carte Phytogéog. Ser. B 7: 99-109.

De Foucault B., 1979. Elements pour une thèorie statistique du tableau phytosociologique homogéne. Vegetatio 40: 163-174.

Dołgopoł M., 1971. Rozkład frekwencji gatunków w zależności od wielkości powierzchni. Master's thesis Warsaw University.

Géhu J. M., Géhu J., 1969. Les assotiations végétales des dunes mobiles et des bordures de la plage de la côte atlantique francaise. Vegetatio 18: 12-166.

Géhu J. M., Géhu J., 1975a. Les fourrés des sables litoraux du Sud-Ouest de la France. Beitr. Naturk. Forsch. 34: 79-94.

Géhu J. M., Géhu J., 1975b. Les fourrés à *Erica scoparia* et *Frangula alnus* d'Aquitaine (*Scopario-Franguletum alnae* J. M. et J. Géhu 1973). Doc. Phytosoc. 14: 117-120.

Géhu J. M., 1976. Approche phytosociologique synthétique de la végétation des vases salées du littoral atlantique francais (synsystématique et synchorologie). Coll. Phytosoc. 4: 395-462.

Greig-Smith P., 1952. The use of random and contiguos quadrats in the study of the structure of plant communities. Ann. Bot. 16: 293-316.

Guinochet M., 1954. Sur les fondements statistiques de la phytosociologie et quelgues-unes de leurs conséquences. Veröff. Geobot. Inst. 29: 41-67.

Guinochet M., 1955. Logique et dynamique du peuplement végétal. Masson, Paris.

Guinochet M., 1957. Les associations végétales. Stientia 92: 179-184.

Kwiatkowska A. J., 1972. Analiza homogeniczności runa fitocenozy przy zastosowaniu nieparametrycznego testu serii. Phytocoenosis 1: 5-37.

Matuszkiewicz W., Wydrzycka U., 1972. Liczba gatunków jako funkcja powierzchni i problem reprezentatywnej powierzchni fiotocenozy. Phytocoenosis 1: 95-120.

Pearson K., 1902. On the systematic fitting of curves to observation and measurement. Biometrica 1: 265-303 and 2: 1-23.

Raunkiaer C., 1918. Rechercher statistiques sur les formations végétales. Biol. Meddeldser.

Yule G. U., Kendall M. G., 1966. Wstęp do teorii statystyki. PWN, Warszawa.

Rozkład frekwencji gatunków jako funkcja wielkości powierzchni

Streszczenie

Badania nad zależnością typu rozkładu frekwencji gatunków od wielkości powierzchni przeprowadzono w runie dwóch fitocenoz leśnych, dąbrowie i grądzie. Ze względu na porównawczy cel pracy w zbiorze i opracowaniu materiału przyjęto schemat postępowania za Dagnelie (1962). Dla uzyskanych rozkładów frekwencji obliczano średnią arytmetyczną ($\bar{x}$), wariancję (S^2), współczynniki skośności i kurtozy rozkładu: β_1, γ_1, β_2 i K Pearsona.

Wyniki analizy zależności empirycznego rozkładu frekwencji od wielkości powierzchni przedstawiają rysunki 1 i 2. W obu przypadkach histogramy frekwencji są początkowo typu L i przekształcają się stopniowo wraz ze zwiększaniem się wielkości powierzchni próbnej w typ U. Empiryczny rozkład frekwencji jest symetryczny w sensie statystycznym, gdy wartości współczynników β_1 i K są równe zeru (rys. 6 i 7) a reprezentuje symetrię typu U gdy wartość β_2 jest mniejsza od 1,8 (rys. 8). W dąbrowie sytuacja taka zachodzi dla powierzchni próbnej równej 32 m^2, a w grądzie — 64 m^2. Zgodnie z przyjętym przez Dagnelie kryterium jednorodności spełniony jest tym samym warunek jednorodności florystycznej badanych płatów. Z przeprowadzonej analizy wynika, że typ rozkładu frekwencji, wartość charakterystyk statystycznych rozkładu i stopień homogeniczności zależą od wielkości powierzchni.

Wielkość powierzchni, na której fitocenoza jest jednorodna z uwagi na przyjęte kryterium może być uważana za optymalną wielkość powierzchni próbnej, reprezentatywnej florystycznie dla danej fitocenozy. Można ją łatwo wyznaczyć z wykresu zależności współczynnika γ_1 (rys. 4) od wielkości powierzchni.