

Grundzüge der Statistischen Ökologie

D. Pfeifer

H.-P. Bäumler

U. Schleier

V. de Valk



Inhalt

Was Ökologie ist und was nicht	1
Zur Rolle der Mathematik in der Ökologie	3
1. Statistische Aspekte von Freilandexperimenten	6
Schätzung von Populationsanteilen	6
Capture–Recapture Experimente	11
Transekt–Verfahren	23
Abstands–Verfahren	30
Versuchsplanung	38
2. Analyse und Vergleich von Lebensgemeinschaften	48
Erfassung von Arten	49
Diversitätsindices	67
Literatur	85

Was Ökologie ist und was nicht

Das Wort "Ökologie" wurde erstmals von Ernst Haeckel im Jahre 1869 benutzt. Frei nach Haeckel können wir Ökologie als das wissenschaftliche Studium der Wechselbeziehungen zwischen Organismen und ihrer Umwelt definieren. Das Wort leitet sich vom griechischen Wort *oikos*, "Heimat" – "Wohnort", ab. Man kann Ökologie also als das Studium des "Familienlebens" lebender Organismen betrachten. Eine bessere, weniger vage Definition wurde von Krebs (1972) vorgeschlagen: "Ökologie ist die wissenschaftliche Untersuchung jener Wechselbeziehungen, welche die Verbreitung und Häufigkeit von Organismen bestimmen." Die Definition hat den Verdienst, daß sie den genauen Gegenstand der Ökologie erfaßt: die Verbreitung und Häufigkeit von Organismen – wo Organismen vorkommen, wieviele es von ihnen gibt, und was sie tun.

...

Was den Gegenstand der Ökologie betrifft, "die Verbreitung und Häufigkeit von Organismen", ist er erfreulich prägnant gefaßt. Wir müssen ihn jedoch erweitern. Ökologie befaßt sich mit drei Stufen: mit einzelnen *Organismen*, der *Population* (aus Individuen einer Art bestehend) und der *Lebensgemeinschaft* (aus einer mehr oder weniger großen Anzahl von Populationen bestehend).

Auf der Stufe des Organismus befaßt sich die Ökologie damit, wie Individuen von ihrer biotischen und abiotischen Umwelt beeinflußt werden, und wie sie diese beeinflussen. Auf der Stufe der Population untersucht die Ökologie das Vorhandensein oder das Fehlen bestimmter Arten, ihre Häufigkeit oder Seltenheit und die Trends und Fluktuationen ihrer Anzahl. ... Die Ökologie von Lebensgemeinschaften behandelt deren *Zusammensetzung* oder *Struktur*, und interessiert sich für den Fluß von Energie, Nährstoffen und anderen chemischen Stoffen durch die Lebensgemeinschaften, also mit ihrem *Funktionieren*.

...

Es ist vielleicht weniger selbstverständlich, daß Ökologen sich oft auch für Laborsysteme und **mathematische Modelle** interessieren. Diese waren für die Entwicklung der Ökologie jedoch von entscheidender Bedeutung und werden es auch in der Zukunft sein. ... Ein wesentliches Anliegen der Wissenschaft ist es zu vereinfachen und auf diese Weise zu versuchen, die Komplexität der realen Welt leichter zu verstehen. Daraus folgt, daß wir an der realen Welt interessiert sind, und der Wert von Modellen und einfachen Laboruntersuchungen muß stets daran gemessen werden, inwieweit sie die Funktion von natürlichen Systemen erhellen. Sie sind Mittel zum Zweck, aber nicht der Zweck selbst.

[aus Begon/Harper/Townsend (1991)]

Die Betonung des naturwissenschaftlichen Charakters der Ökologie widerspricht einer Erwartungshaltung, die in weiten Teilen unserer Gesellschaft verbreitet ist. Es herrscht vielfach die Meinung, die Ökologie könne der Grundlegung einer neuen Ethik oder einer neuen politischen Doktrin dienen. Derartige Erwartungen stecken in Formulierungen wie "Ökologie und Ökonomie versöhnen", "ökosoziale Marktwirtschaft", "Ökosozialismus" etc. Solche Forderungen an eine Naturwissenschaft widersprechen sowohl deren Wesen als auch dem Wesen einer demokratischen Gesellschaft.

Naturwissenschaften können Theorien darüber anbieten, wie die Natur funktioniert. Sie können jedoch nicht angeben, welcher Zustand der Natur ein erhaltenswürdiger Wert oder ein erstrebenswertes Ziel menschlichen Handelns ist. Die Festsetzung solcher Werte und Ziele ist keine Aufgabe der Wissenschaft, sondern eine Aufgabe des demokratischen Entscheidungsprozesses. Weder Ökologen noch irgendwelche anderen Experten haben in diesem Entscheidungsprozeß größere Rechte als andere Menschen. Die Aufgabe des Experten besteht allenfalls darin, Wege zum Erreichen von Zielen aufzuzeigen und Konflikte zwischen verschiedenen, für sich genommen durchaus berechtigten Zielen frühzeitig zu erkennen.

...

Die ungeheuere Popularität, die Umweltfragen inzwischen erreicht haben, ist sehr positiv zu bewerten, hat aber zu einer starken Verschiebung der Gewichte zur angewandten Ökologie geführt. Deren Zielsetzung ist nicht der Aufbau einer generellen Theorie, sondern die Optimierung der Lebensbedingungen des Menschen, und diese ist nicht immer mit wissenschaftlichen Kriterien zu begründen. ... Das Verhältnis zwischen Ökologie und Umweltwissenschaft kann man vielleicht mit dem Verhältnis zwischen Physik und Ingenieurwissenschaften vergleichen. Die Physik liefert die Grundlagen, die Ingenieure machen daraus Produkte, die dem Menschen nutzen sollen. Während aber Physik und Ingenieurwissenschaften relativ klar abgegrenzt sind, ist das bei Ökologie und Umweltwissenschaft nicht der Fall. Nur ein Bruchteil der Fördermittel, die für "Ökologische Forschung" aufgewendet werden, geht in die ökologische Grundlagenforschung. Der größte Teil wird für die Bewältigung aktueller Probleme und für "Monitoring" aufgewendet. Solche Programme sind dringend notwendig und sollten eher noch verstärkt werden, sie bringen aber wenig wissenschaftlichen Fortschritt, da sie in der Regel darin bestehen, bekannte Konzepte und Methoden anzuwenden. Die verstärkten Aufwendungen haben deshalb bisher leider nicht zu einem entsprechenden Zuwachs an ökologischer Theorie geführt.

...

Es wäre wünschenswert, den Begriff "Ökologie" wieder für das zu verwenden, was ursprünglich bezeichnete, eine Teildisziplin der Biologie mit eigenen Methoden und einem eigenen Theoriengebäude, das anderen Teildisziplinen gleichwertig ist, weder eine "integrierende" Wissenschaft noch eine Heilslehre.

[aus Lampert/Sommer (1993)]

Zur Rolle der Mathematik in der Ökologie

The fact that ecology is essentially a mathematical subject is becoming ever more widely accepted. Ecologists everywhere are attempting to formulate and solve their problems by mathematical reasoning, using whatever mathematical knowledge they have acquired, usually in undergraduate courses or private study.

...

Ecological model building takes many different forms, depending on the purpose for which it is done. At one extreme are so-called "explanatory" models or systems models, whose behavior is thought to duplicate, at least approximately, the true behavior of the populations being modeled. . . . Properly applied, they do contribute to our understanding of ecological processes. It is a mistake to describe these models as "useful" if they fit actual observations (and, by extension, useless if they do not). Such a notion involves a misunderstanding of the purpose of theoretical models. A poor fit is often more informative than a good one. A good fit may imply no more than that the observations and the model reproducing them are not capable of discriminating among competing theories on ecosystem functioning. A poor fit usually permits rejection of an untenable theory and to that extent at least yields a gain in knowledge. Even a small gain is better than none.

In contrast to models designed only to *explain* the behavior of populations are models designed to *predict* their behavior in the future. The great importance of such models in all branches of applied ecology is obvious. They strive to answer the question: how will a given ecosystem continue to develop if left to itself, and how will it react if interfered with?

...

The kinds of ecosystems whose structure and function theoretical ecologists seek to understand cover a very wide range in terms of size and complexity. At one end of the spectrum are laboratory microcosms in which a few species of small, short-lived, fast-breeding, active animals (for example, beetles, fruit flies, blowflies, *Daphnia*¹, *Paramecia*²) live and interact for many generations under controlled conditions in a container. At the other end are large natural systems of baffling complexity; systems in which thousands or ten thousands of species live together under continually changing conditions and in spatially heterogeneous environments. The more complex the ecosystem an ecologist studies, the less useful are simple explanatory "system models" as aids to understanding and the more useful is a "statistical" approach. The contrast is worth examining. A "modeler's" conjectures are about abstractions; he envisages possible interactions among a few idealized species-populations and deduces their consequent behavior; only then (if at all!) is the mental construct compared with reality. A "statistician's" conjectures are about concrete, observable

¹ Wasserflöhe

² Wimpertierchen

things; contemplating real ecosystems, which initially furnish an unstructured welter of impressions, he strives to single out those observable phenomena that serve as clues to underlying processes. It is detective work. It entails discrimination, in a situation where examination of all the facts is not feasible, between facts that are, and facts that are not, relevant to the solution of particular problems. The problems themselves are usually prompted by observation.

It is worth remarking that ecosystems whose understanding demands the "statistical" approach are of greater concern to ecologists than are those that lend themselves to the "modeling" approach.

...

The third approach to mathematical-plus-statistical ecology is closely related to the second. It consists in **processing large bodies of observational data** in such a way that interesting regularities, hitherto buried from sight, become apparent. These regularities can inspire the building of hypotheses, and the hypotheses can then be tested by appeal to other bodies of the data. Thus approaches 2 and 3 are both **statistical**: both require that evidence from real ecosystems be sifted and interpreted. The difference is that testable hypotheses are derived from raw data in approach 2 and from processed data in approach 3. The processing may be done in numerous ways. Methods of handling multivariate statistical data are constantly being improved as computers increase in sophistication; new methods are being tested in ecological contexts as soon as they are developed. The aim in all cases is to simplify and clarify unwieldy masses of "noisy" data so that their underlying structure is revealed. Only then, perhaps, can it be explained.

[aus Pielou (1977)]

However, neither the application of mere statistical methods nor the sophisticated analysis of mathematically interesting models alone suffice to put forward a deeper understanding of biological processes. A synthesis of both statistical methods and mathematical models is most promising in our opinion. This is also related to the acceptance of models by experimentalists: nothing is more convincing than a model trajectory running through the data points.

[aus Richter/Söndgerath (1990)]

Change from a qualitative to a quantitative approach is characteristic of the development of any branch of science. As some understanding is achieved of the broader aspects of the phenomena, interest naturally turns to the finer detail of structure or behaviour, in which the observable differences are smaller and can only be appreciated in terms of measurement. It is not surprising that a quantitative outlook has been attained earlier in most branches of physical science than in biological science. Perhaps the greatest single difference in methodology between the physical and biological sciences is that in the former it is generally possible to isolate one variable at a time for study, whereas in the latter this is rarely possible. Thus, in the physical sciences broad outlines of phenomena are more readily seen from a relatively simple programme of qualitative investigation, and the way cleared for the more exact quantitative approach. In biology not only is it rarely possible to isolate variables for study, but the subjects of investigation are themselves commonly so complex that they are difficult to measure.

...

The general impossibility of controlling 'unwanted' variables in biology leads to a much greater degree of error variability in measurement than in the physical sciences. In the physical sciences differences among replicate measurements are generally attributable to deficiencies of technique, whereas in biological observations differences may be due not only to these deficiencies, but also, and commonly to a much greater extent, to fluctuation in variables not under investigation and assumed to be constant. Put another way, it is very much more difficult to obtain truly replicate samples in biological measurements than in physical measurements. If measurements are made in two or more contexts with the object of determining if there is any difference in the variable measured, the means may be different but the ranges of individual measurements may overlap. Thus the problem arises whether an observed difference is significant or not, i.e. whether it reflects any real difference between the two groups sampled, or is due to chance. In the physical sciences the immediate reaction is to improve the technique to obtain more accurate measurements. . . . In biology, however, there is often little scope for improvement of technique, and the biologist is therefore forced to turn for help in judging significance of difference to the techniques of **statistical analysis**. These are based on probability theory and permit determination of the probability of observed differences arising by chance in different samples of the same population. Thus arises the apparent paradox that while the physical sciences make much greater use of a quantitative approach than the biological sciences, they are much less dependent on the techniques of statistical analysis for the interpretation of their quantitative data.

[aus Greig-Smith (1983)]

1. Planung, Durchführung und Auswertung von Freilandexperimenten

Die Wissenschaft schafft Wissen, und Wissen besteht aus Antworten auf Fragen. In den Naturwissenschaften drehen sich diese Fragen um Dinge, die – wenigstens im Prinzip – zählbar oder meßbar sind. Jede gezielte wissenschaftliche Untersuchung beginnt mit einer Frage. ... Aus der gestellten Frage ergibt sich, welche Information Sie brauchen, um sie beantworten zu können. Fast immer sind dazu viele Beobachtungen oder Messungen nötig. Deshalb spricht man auch von "quantitativen" Untersuchungen. In einem nächsten Schritt müssen Sie die notwendigen Daten sammeln und analysieren. Und wenn nichts Unvorhergesehenes dazwischen kommt, müßte das Ergebnis dann Schlußfolgerungen erlauben, die Ihre Ausgangsfrage beantworten.

[aus Lamprecht (1992)]

Schätzung von Populationsanteilen

Eine der einfachsten Fragestellungen im Zusammenhang mit Zählungen von Pflanzen und Tieren ist die Schätzung von *Populationsanteilen*, z.B. dem Anteil juveniler / adulter Individuen in einer Population. Dabei sind häufig weder die Größe N der Gesamtpopulation noch der Umfang K der interessierenden Teilpopulation bekannt. Die gewünschte Größe $p = \frac{K}{N}$ kann daher in der Regel nur durch die Auswertung einer *Stichprobe* vom Umfang $n < N$ geschätzt werden. Wenn die Zufallsgröße X die Anzahl der in der Stichprobe vorkommenden Individuen der interessierenden Teilpopulation bezeichnet, ergibt sich als deren Verteilung (*hypergeometrische Verteilung*):

$$P(X = k) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}, \quad k = 0, \dots, m := \min\{n, K\}.$$

Als Schätzung für den gewünschten Anteil wählt man in der Regel die Größe $\hat{p} := \frac{X}{n}$, d.h. den in der Stichprobe enthaltenden Anteil der interessierenden Teilpopulation. Dieser Schätzer ist *erwartungstreu*, d.h. es gilt $E(\hat{p}) = p$: die auf \hat{p} beruhende Schätzung des "wahren" Anteils p enthält keine systematischen Fehler. Die zugehörige Rechnung ist

$$\begin{aligned} E(X) &= \sum_{k=0}^m k \cdot P(X = k) = \frac{1}{\binom{N}{n}} \sum_{k=0}^m k \binom{K}{k} \binom{N-K}{n-k} \\ &= \frac{K}{\binom{N}{n}} \sum_{k=1}^m \binom{K-1}{k-1} \binom{(N-1)-(K-1)}{(n-1)-(k-1)} = K \cdot \frac{\binom{N-1}{n-1}}{\binom{N}{n}} = K \frac{n}{N} = np, \end{aligned}$$

also $E(\hat{p}) = p$. Dabei wurde die folgende Beziehung für Binomialkoeffizienten benutzt:

$$k \binom{K}{k} = k \frac{K!}{k!(K-k)!} = \frac{K!}{(k-1)!(K-k)!} = K \binom{K-1}{k-1}, \quad k = 1, \dots, K.$$

Die Genauigkeit dieses Schätzers kann vermöge der zugehörigen Varianz bestimmt werden; hierfür gilt:

$$\sigma^2(\hat{p}) = \frac{p(1-p)}{n} \frac{N-n}{N-1} \leq \frac{1}{4n}.$$

Es ist nämlich ähnlich wie zuvor

$$\begin{aligned} E(X(X-1)) &= \sum_{k=0}^m k(k-1) \cdot P(X=k) = \frac{1}{\binom{N}{n}} \sum_{k=0}^m k(k-1) \binom{K}{k} \binom{N-K}{n-k} \\ &= K(K-1) \frac{\binom{N-2}{n-2}}{\binom{N}{n}} = \frac{K(K-1)n(n-1)}{N(N-1)} \end{aligned}$$

und somit

$$E(X^2) = E(X(X-1)) + E(X) = \frac{nK}{N} \left(1 + \frac{K-1}{N-1}(n-1) \right),$$

woraus

$$\sigma^2(X) = E(X^2) - [E(X)]^2 = pn \left(1 - \frac{K}{N} \right) \frac{N-n}{N-1} = np(1-p) \frac{N-n}{N-1}$$

und damit

$$\sigma^2(\hat{p}) = \frac{1}{n^2} \sigma^2(X) = \frac{p(1-p)}{n} \frac{N-n}{N-1}$$

folgt (siehe hierzu auch Richter/Söndgerath (1990), S. 13f.). Die restliche Abschätzung ergibt sich aus der Tatsache, daß die Abbildung $p \mapsto p(1-p)$ für $p \in [0, 1]$ maximal bei $p = 1/2$ ist mit Wert $1/4$.

Eine einfache Abschätzung der Präzision der Schätzung erhält man z.B. durch die *Tschebyscheff-Ungleichung*:

$$P(|\hat{p} - p| > \varepsilon) \leq \frac{\sigma^2(\hat{p})}{\varepsilon^2} \leq \frac{1}{4n\varepsilon^2}, \quad \varepsilon > 0,$$

die unabhängig von den unbekanntem Größen K und N ist. Will man beispielsweise mit mindestens 90%iger Sicherheit den Anteil p mit einem maximalen Fehler von 5% vorhersagen, so ist (bei vorsichtiger Herangehensweise) $\varepsilon = 0,05$ und n so zu wählen, daß

$$\frac{1}{4n\varepsilon^2} = \frac{100}{n} \leq \frac{1}{10}, \quad \text{also } n \geq 1000$$

ausfällt. Ist die Gesamtgröße N der Population (zumindest näherungsweise) bekannt, kann unter Verwendung der günstigeren Abschätzung

$$P(|\hat{p} - p| > \varepsilon) \leq \frac{\sigma^2(\hat{p})}{\varepsilon^2} \leq \frac{1}{4n\varepsilon^2} \frac{N-n}{N-1}, \quad \varepsilon > 0,$$

der Stichprobenaufwand bei gleicher Genauigkeit verringert werden. Bezeichnet α die tolerierte Fehlerwahrscheinlichkeit für die Vorhersage von p , so folgt durch Umstellen der Ungleichung nämlich

$$n \geq \frac{N}{1 + 4\varepsilon^2(N-1)\alpha}.$$

Legt man für das genannte Beispiel etwa eine Gesamtpopulation von $N = 1000$ zugrunde, so reduziert sich bei gleicher Präzision der berechnete Stichprobenaufwand von $n \geq 1000$ auf $n \geq 501$, also etwa auf die Hälfte.

Bei konkreten Anwendungen kommt es gelegentlich vor, daß zur Schätzung von Populationsanteilen an r räumlich voneinander getrennten Standorten Stichproben vom Umfang n_1, \dots, n_r erhoben werden. In der Ornithologie ist das z.B. der Fall, wenn die gesamte Population sich in verschiedenen Brutgebieten über die untersuchte Fläche verteilt. Man erhält dann zunächst r einzelne Schätzer $\hat{p}_1, \dots, \hat{p}_r$ für den gesamten Populationsanteil p , die – zufallsbedingt – um den wahren Wert p schwanken. Es ist naheliegend, die einzelnen Schätzungen zu einer einzigen Schätzung zusammensetzen, z.B. durch eine geeignete Mittelbildung. In der Statistik betrachtet man oft *lineare* Schätzer, d.h. man macht einen Ansatz

$$\hat{p} = \sum_{i=1}^r a_i \hat{p}_i$$

mit geeigneten Gewichten a_1, \dots, a_r . Unter der Voraussetzung der *Erwartungstreue* an den Gesamtschätzer \hat{p} müssen diese Gewichte dann die Nebenbedingung $\sum_{i=1}^r a_i = 1$ erfüllen. Wählt man als Optimierungskriterium für die Präzision der Gesamtschätzung deren Varianz, so ergibt sich das Lagrange-Problem

$$\min! \quad \sigma^2(\hat{p}) = \sigma^2\left(\sum_{i=1}^r a_i \hat{p}_i\right) \quad \text{unter der Nebenbedingung} \quad \sum_{i=1}^r a_i = 1.$$

Mathematisch läßt sich dieses Problem übersichtlicher in vektorieller Form behandeln, wenn man die zu optimierenden Gewichte zu dem Spaltenvektor $\mathbf{a} = (a_1, \dots, a_r)^{tr}$ zusammenfaßt. Bezeichnet Σ die Varianz-Kovarianzmatrix der zugehörigen Einzelschätzungen $\hat{\mathbf{p}} = (\hat{p}_1, \dots, \hat{p}_r)^{tr}$ innerhalb der Teilstichproben, also $\Sigma = (\sigma_{ij})$ mit

$$\sigma_{ij} = \text{Kov}(\hat{p}_i, \hat{p}_j) = E(\hat{p}_i \hat{p}_j) - E(\hat{p}_i)E(\hat{p}_j), \quad 1 \leq i, j \leq r,$$

so reduziert sich das gegebene Lagrange–Problem auf die Minimierung einer quadratischen Form:

$$\min! \mathbf{a}^{tr} \Sigma \mathbf{a} \quad \text{unter} \quad \mathbf{a}^{tr} \mathbf{1} = 1,$$

wobei der Spaltenvektor $\mathbf{1} = (1, \dots, 1)^{tr}$ den Vektor bezeichnet, dessen Komponenten sämtlich 1 sind. Die allgemeine Lösung des Lagrange–Problems erhält man aus der Minimierung der Lagrange–Funktion

$$L(\mathbf{a}, \lambda) := \mathbf{a}^{tr} \Sigma \mathbf{a} - \lambda(\mathbf{a}^{tr} \mathbf{1} - 1),$$

mit der notwendigen Bedingung¹

$$\nabla_{\mathbf{a}} L(\mathbf{a}, \lambda) = 2\Sigma \mathbf{a} - \lambda \mathbf{1} = \mathbf{0} \quad (\text{Nullvektor}).$$

Sofern die Inverse Σ^{-1} existiert, hat also \mathbf{a} die Form $\mathbf{a} = c \Sigma^{-1} \mathbf{1}$ mit einer geeigneten Konstanten $c = \lambda/2$. Durch Einsetzen in die Nebenbedingung ergibt sich

$$c = \frac{1}{\mathbf{1}^{tr} \Sigma^{-1} \mathbf{1}}$$

und damit als minimale Varianz

$$\sigma^2(\hat{p}) = \mathbf{a}^{tr} \Sigma \mathbf{a} = c^2 \mathbf{1}^{tr} \Sigma^{-1} \Sigma \Sigma^{-1} \mathbf{1} = c = \frac{1}{\mathbf{1}^{tr} \Sigma^{-1} \mathbf{1}}.$$

Für den hier betrachteten konkreten Anwendungsfall benötigen wir zunächst die Kovarianzen der Zufallsgrößen X_1, \dots, X_r der in den Teilstichproben enthaltenen interessierenden Individuenzahlen. Hier gilt für $i \neq j$:

$$\text{Kov}(X_i X_j) = -\frac{p(1-p)}{N-1} n_i n_j, \quad 1 \leq i, j \leq r.$$

Ähnliche Überlegungen wie oben zeigen nämlich

$$P(X_i = k_i, X_j = k_j) = \frac{\binom{K}{k_i, k_j} \binom{N-K}{n_i - k_i, n_j - k_j}}{\binom{N}{n_i, n_j}},$$

wobei $k_i = 0, \dots, n_i$, $k_j = 0, \dots, n_j$ und $k_i + k_j \leq K$. Dabei ist der Multinomialkoeffizient $\binom{K}{k_i, k_j}$ erklärt als

$$\binom{K}{k_i, k_j} = \frac{K!}{k_i! k_j! (K - k_i - k_j)!}.$$

¹ $\nabla_{\mathbf{a}} L = \left(\frac{\partial L}{\partial a_1}, \dots, \frac{\partial L}{\partial a_r} \right)^{tr}$ bezeichnet den Gradienten von L nach \mathbf{a} .

Für die Kovarianzen erhält man somit zunächst

$$\begin{aligned}
 E(X_i X_j) &= \sum_{\substack{1 \leq k_i \leq n_i, 1 \leq k_j \leq n_j \\ k_i + k_j \leq K}} k_i k_j P(X_i = k_i, X_j = k_j) \\
 &= \frac{1}{\binom{N}{n_i, n_j}} \sum_{\substack{1 \leq k_i \leq n_i, 1 \leq k_j \leq n_j \\ k_i + k_j \leq K}} k_i k_j \binom{K}{k_i, k_j} \binom{N-K}{n_i - k_i, n_j - k_j} \\
 &= \frac{K(K-1)}{\binom{N}{n_i, n_j}} \sum_{\substack{1 \leq k_i \leq n_i, 1 \leq k_j \leq n_j \\ k_i + k_j \leq K}} \binom{K-2}{k_i - 1, k_j - 1} \binom{(N-2) - (K-2)}{(n_i - 1) - (k_i - 1), (n_j - 1) - (k_j - 1)} \\
 &= K(K-1) \frac{\binom{N-2}{n_i - 1, n_j - 1}}{\binom{N}{n_i, n_j}} = \frac{K(K-1)}{N(N-1)} n_i n_j
 \end{aligned}$$

und damit

$$\text{Kov}(X_i, X_j) = \left(\frac{K(K-1)}{N(N-1)} - \frac{K^2}{N^2} \right) n_i n_j = -\frac{p(1-p)}{N-1} n_i n_j$$

bzw.

$$\sigma_{ij} = \frac{1}{n_i n_j} \text{Kov}(X_i, X_j) = -\frac{p(1-p)}{N-1}.$$

Die Matrix Σ erhält damit die Form

$$\Sigma = \frac{p(1-p)}{N-1} \begin{pmatrix} \frac{N}{n_1} - 1 & -1 & -1 & \dots & -1 \\ -1 & \frac{N}{n_2} - 1 & -1 & \dots & -1 \\ \vdots & \vdots & \vdots & & \vdots \\ -1 & -1 & -1 & \dots & \frac{N}{n_r} - 1 \end{pmatrix},$$

woraus sich für die Lösung des Lagrange-Problems ergibt:

$$a_i = \frac{n_i}{n_1 + \dots + n_r}, \quad i = 1, \dots, r.$$

Dies sieht man direkt durch Berechnung des Vektors $\Sigma \mathbf{a}$, der mit dieser Wahl wie verlangt lauter gleiche Komponenten besitzt. Die optimale Wahl der Gewichte a_1, \dots, a_r hängt also insbesondere nicht von den i.a. unbekanntem Parametern N und K ab und ist direkt proportional zu den Größen der Teilstichproben n_1, \dots, n_r . Anschaulich stimmt der Schätzer \hat{p} also mit demjenigen Schätzer überein, der sich durch Poolen der Teilstichproben ergeben würde:

$$\hat{p} = \frac{X_1 + \dots + X_r}{n_1 + \dots + n_r}$$

und besitzt daher wie oben die (minimale) Varianz

$$\sigma^2(\hat{p}) = \frac{p(1-p)}{n_1 + \dots + n_r} \frac{N - (n_1 + \dots + n_r)}{N-1}.$$

Capture–Recapture Experimente

Es gibt zwei unterschiedliche Qualitäten von Häufigkeitsangaben: relative und absolute. ... Hängt das Fangergebnis von der Aktivität der Tiere ab, werten wir die angetroffene Häufigkeit als "Aktivitätsdichte" im Unterschied zu Maßen, die Tierzahlen pro Flächen- oder Volumeneinheit angeben ("Individuendichte", Abundanz). Für die Bearbeitung ökologischer Probleme sollte immer eine Abundanzbestimmung angestrebt werden: an Abundanzschwankungen einzelner Populationen werden z.B. Räuber–Beute–Beziehungen, interspezifische Konkurrenz oder Stabilität von Ökosystemen studiert und für Schätzungen der Überlebenswahrscheinlichkeit einer Art (Naturschutz) oder des Energieflusses im Ökosystem ist die Kenntnis der Populationsgrößen ebenfalls erforderlich.

...

Zur problemlosen Anwendung von Berechnungsmethoden einer Wiederfang–Analyse (mark–recapture, capture–recapture) sind eine Reihe von Voraussetzungen zu überdenken:

- Die Tiere dürfen durch die Markierung in ihrer Wiederfangbarkeit nicht beeinträchtigt werden. Die Markierungen müssen die gesamte Fangperiode über ablesbar bleiben (Laborkontrolle).
- Die markiert–freigelassenen Tiere müssen sich vollständig in die Population mischen. Alle Tiere müssen gleich fangbar sein.

...

Diese Bedingungen sind [für die Erfassung der Fauna in einer Wiese] hinreichend gut erfüllt bei Laufkäfern (*Carabidae*), Sandlaufkäfern (*Cicindelidae*), Wolfspinnen (*Lycosidae*), Zikaden (*Cicadina*), Feldheuschrecken (*Caelifera*), Kleinlibellen (*Zygoptera*) in Feucht– oder Uferwiesen, ausgewählte Fliegengruppen (*Brachycera*) und Schmetterlingen (*Lepidoptera*), Kleinsäugetern (v.a. *Arvicolidae*¹ und *Muridae*²) und Vögeln (*Passeriformes*³), in Sonderfällen auch für Gehäuseschnecken.

...

Eine Laufkäferpopulation kann man untersuchen, indem man die Tiere mit einem Gitternetz an Bodenfallen lebend fängt, im Labor oder Freiland mit einem Thermocauter auf den Flügeldecken markiert und wieder freilässt. ... Laufspinnen sind schwieriger zu markieren. Damit nach dem Fang der Tiere in Bodenfallen beim Festhalten keine Beine verloren gehen oder schon im Feld die Tiere besser mit der Hand eingesammelt werden können, besprüht man die Tiere kurz mit einem Vereisungsmittel. Eine Anästhesierung der Spinnen mit Kohlendioxyd ist umständlicher, funktioniert aber auch in dichter Vegetation wie in Grasbüscheln. ... Als Farbmarkierung eignen sich Nagellacke, Modellfarben, Bastelfarben, die wie Dispersionsfarben wasserunlöslich werden, oder kleine Dosen von Autolacken. ... Zur Kontrolle, ob

¹ Wühlmäuse

² Mäuse, Ratten

³ Sperlingsvögel

die Markierung dem Tier schadet, hält man sich markierte Tiere vorübergehend in einem Terrarium zumindest für die Dauer der Freilanduntersuchung.

...

Mit individuell markierten Tieren eröffnet sich durch Wiederfang bzw. Wiederbeobachtung neben der Dichtebestimmung ein weites Feld von Möglichkeiten, Aufenthalts- und Wanderungskontrollen durchzuführen. Studien des Bewegungsmusters sind gerade für Tiere in der vom Menschen immer wieder veränderten Kulturlandschaft von größtem Interesse.

[aus Mühlenberg (1993)]

Mathematisch ist das Problem der Auswertung von Wiederfangergebnissen markierter Tiere mit der Schätzung von Populationsanteilen eng verwandt. Im Unterschied zur vorigen Problemstellung legt der Wissenschaftler allerdings hier die Größe K , nämlich die Anzahl der insgesamt markierten Tiere, selbst fest. Der unbekannte Populationsumfang N ist dann ähnlich wie zuvor auf der Basis der Zufallsgröße X – das ist hier die Anzahl der beim Wiederfang in einer Stichprobe vom Umfang n gezählten markierten Tiere – zu schätzen. Aufgrund der schon hergeleiteten Beziehung $E(X) = \frac{nK}{N}$ liegt es nahe, dafür den Schätzer

$$\hat{N} = \frac{nK}{X} \quad (\text{Petersen- bzw. Lincoln-Index})$$

zu verwenden. Die Zufallsgröße X kann allerdings den Wert Null annehmen (wenn in der Stichprobe keine markierten Tiere vorhanden sind). Der Schätzer \hat{N} ist in diesem Fall also gar nicht definiert! Zur Vermeidung dieser Schwierigkeiten ist es günstiger, einen Schätzer der Form

$$\hat{N} = \frac{c}{X+1} + d$$

mit geeigneten Konstanten c, d zu betrachten, der möglichst auch erwartungstreu ist. Eine ähnliche Rechnung wie oben ergibt ($m = \min\{n, K\}$):

$$\begin{aligned} E\left(\frac{1}{X+1}\right) &= \sum_{k=0}^m \frac{1}{k+1} \cdot P(X=k) = \frac{1}{\binom{N}{n}} \sum_{k=0}^m \frac{1}{k+1} \binom{K}{k} \binom{N-K}{n-k} \\ &= \frac{1}{(K+1)\binom{N}{n}} \sum_{k=0}^m \binom{K+1}{k+1} \binom{(N+1)-(K+1)}{(n+1)-(k+1)} \\ &= \frac{1}{K+1} \cdot \frac{\binom{N+1}{n+1} - \binom{N-K}{n+1}}{\binom{N}{n}} = \frac{N+1}{(n+1)(K+1)} \left[1 - \frac{\binom{N-K}{n+1}}{\binom{N+1}{n+1}} \right]. \end{aligned}$$

Wir betrachten zunächst den Fall $N - K < n$, d.h. der Stichprobenumfang des Wiederfangs ist größer als der nicht-markierte Teil der Gesamtpopulation. Der obige Ausdruck vereinfacht sich dann zu

$$E\left(\frac{1}{X+1}\right) = \frac{N+1}{(n+1)(K+1)},$$

d.h. der Schätzer

$$\hat{N} = \frac{(n+1)(K+1)}{X+1} - 1 \quad (\text{modifizierter Petersen- bzw. Lincoln-Index})$$

ist in diesem Fall ein erwartungstreuer Schätzer für N mit $c = (n+1)(K+1)$, $d = -1$. Das Verfahren ist allerdings praxisuntauglich, da der Umfang des Wiederfangs hier zahlenmäßig in der Größenordnung der Gesamtpopulation liegen muß. In der realistischeren Situation $n \leq N-K$ ist der Schätzer $\hat{N} = \frac{(n+1)(K+1)}{X+1} - 1$ jedoch verzerrt: es ist nämlich

$$E(\hat{N}) = N - (N+1) \frac{\binom{N-K}{n+1}}{\binom{N+1}{n+1}}.$$

Approximativ gilt hier (vgl. Richter/Söndgerath (1990), S. 105):

$$E(\hat{N}) \approx N \left[1 - \exp \left(- \frac{(n+1)(K+1)}{N} \right) \right].$$

Für sehr große Populationen ($N \rightarrow \infty$) erhält man damit

$$\lim_{N \rightarrow \infty} E(\hat{N}) = (n+1)(K+1),$$

d.h. die Schätzung der Populationsgröße ist in diesem Fall nach oben beschränkt. Für eine sinnvolle Anwendung des Petersen-Index sollte daher größenordnungsmäßig $nK > 3N$ sein; in diesem Fall beträgt der maximale relative Bias ungefähr 5%. Ohne wenigstens grobe Vorstellungen hinsichtlich der möglichen Populationsgröße ist die Anwendung dieses Capture-Recapture-Verfahrens also nur mit Einschränkungen zu empfehlen, vor allem, wenn die Zahl der markiert wiedergefangenen Tiere sehr klein ist. Krebs (1985) zitiert in seinem Buch auf S. 163 ein Experiment von Dahl, bei dem 109 norwegische Forellen (*Salmo fario*) markiert wurden. Beim Wiederfang erhielt er 177 Forellen, von denen 57 markiert waren. Der modifizierte Petersen-Schätzer ergibt hier eine Hochrechnung der Forellenpopulation auf $\hat{N} = \frac{(n+1)(K+1)}{X+1} - 1 = \frac{178 \cdot 110}{58} - 1 = 336.59$; der absolute Bias dürfte in diesem Anwendungsbeispiel sehr gering sein (selbst bei einer wahren Population von 3000, also etwa dem 9-fachen Wert der Schätzung, beträgt er weniger als 5!). Die Varianz des modifizierten Petersen-Index läßt sich nicht elementar berechnen; sie ist jedoch für genügend große Werte von nK approximativ gleich

$$\sigma^2(\hat{N}) \approx N \frac{(N-K)(N-n)}{nK}.$$

Die folgenden Tabellen enthalten einige Werte für $\sigma^2(\hat{N})$ und deren Approximation für $N = 100$:

$K = 20$				
$n =$	20	30	40	50
$\sigma^2(\hat{N}) =$	2154.8	1181.5	687.7	433.4
Approximation	1600.0	933.3	600.0	400.0

$K = 30$				
$n =$	20	30	40	50
$\sigma^2(\hat{N}) =$	1181.5	613.2	374.2	242.6
Approximation	933.3	544.4	350.0	233.3

$K = 40$				
$n =$	20	30	40	50
$\sigma^2(\hat{N}) =$	687.7	374.2	233.5	153.1
Approximation	600.0	350.0	225.0	150.0

Die angegebenen Beispielswerte zeigen, daß die Approximation für die Varianz $\sigma^2(\hat{N})$ i.a. nur eine untere Schranke bildet und daher nur für genügend große Werte von nK zur Berechnung von Konfidenzbereichen für die Populationsgröße N geeignet ist.

Die Varianz $\sigma^2(\hat{N})$ kann auch aus der Stichprobe geschätzt werden; Richter/Söndergerath (1990) geben dafür den Schätzer

$$\hat{s}^2 = \frac{(n+1)(K+1)(n-X)(K-X)}{(X+1)^2(X+2)}$$

an, der im Fall $n > N - K$ wieder erwartungstreu ist.

Im Beispiel der norwegischen Forellen liefert dieser Schätzer den Wert $\hat{s}^2 = 615.5$; für eine tatsächliche Population von 337 ergibt sich zum Vergleich eine Varianz von $\sigma^2(\hat{N}) = 642.6$, was mit der berechneten Schätzung noch relativ gut übereinstimmt.

Um der Problematik der geringen Aussagefähigkeit kleiner Anzahlen markierter wiedergefangener Tiere zu entgehen, kann man alternativ auch ein sequentielles Verfahren zur Schätzung der Populationsgröße N anwenden. Dazu werden solange

Tiere wiedergefangen, bis man erstmalig ein markiertes Individuum findet. Bezeichnet Y die Zufallsgröße "Nummer des ersten Fangs eines markierten Tieres", so ist

$$N^* = (K + 1)Y - 1$$

ein erwartungstreuer Schätzer für N , der darüberhinaus sogar ganzzahlige Werte hat. Es gilt nämlich

$$\begin{aligned} P(Y = k) &= \frac{(N - K)(N - K - 1) \cdots (N - K - k + 2)K}{N(N - 1) \cdots (N - k + 1)} \\ &= \frac{K}{N + 1 - k} \prod_{i=0}^{k-2} \left(\frac{N - K - i}{N - i} \right) \\ &= \frac{K}{N + 1 - k} \frac{\binom{N-K}{k-1}}{\binom{N}{k-1}}, \quad k = 1, \dots, N - K + 1 \end{aligned}$$

und damit

$$\begin{aligned} E(N + 1 - Y) &= K \sum_{k=1}^{N-K+1} \frac{\binom{N-K}{k-1}}{\binom{N}{k-1}} \\ &= K \frac{N + 1}{K + 1} \sum_{k=1}^{(N+1)-(K+1)+1} \frac{K + 1}{N + 2 - k} \frac{\binom{(N+1)-(K+1)}{k-1}}{\binom{N+1}{k-1}} = K \frac{N + 1}{K + 1}, \end{aligned}$$

also

$$E(Y) = N + 1 - K \frac{N + 1}{K + 1} = \frac{N + 1}{K + 1},$$

woraus $E(N^*) = N$ unmittelbar folgt. Zur Berechnung der Varianz betrachten wir zunächst analog der obigen Rechnung

$$\begin{aligned} E[(N + 1 - Y)(N + 2 - Y)] &= K \frac{(N + 1)(N + 2)}{K + 2} \sum_{k=1}^{(N+2)-(K+2)+1} \frac{K + 2}{N + 3 - k} \frac{\binom{(N+2)-(K+2)}{k-1}}{\binom{N+2}{k-1}} \\ &= K \frac{(N + 1)(N + 2)}{K + 2}. \end{aligned}$$

Wegen

$$E[(N + 1 - Y)(N + 2 - Y)] = (N + 1)(N + 2) - (2N + 3)E(Y) + E(Y^2)$$

ergibt sich hieraus

$$E(Y^2) = \frac{(2N + 3)(N + 1)}{K + 1} - \frac{2(N + 1)(N + 2)}{K + 2} = \frac{(2(N + 1) - K)(N + 1)}{(K + 1)(K + 2)},$$

also

$$\sigma^2(Y) = \frac{(2(N+1) - K)(N+1)}{(K+1)(K+2)} - \frac{(N+1)^2}{(K+1)^2} = \frac{K(N+1)(N-K)}{(K+1)^2(K+2)}.$$

Als Varianz des Schätzers N^* erhält man also schließlich

$$\sigma^2(N^*) = (K+1)^2 \sigma^2(Y) = \frac{K(N+1)(N-K)}{K+2}.$$

Zur Einordnung dieses Ergebnisses betrachten wir die erwartete Anzahl n von Stichproben, die notwendig sind, um ein markiertes Individuum zu finden. Nach der obigen Rechnung ist dies gerade $n = E(Y) = (N+1)/(K+1) \approx N/K$. Setzen wir diesen Wert formal in die Approximation für $\sigma^2(\hat{N})$ (S. 13) ein, so finden wir

$$\sigma^2(\hat{N}) \approx (N-K)(N - N/K) = \frac{(K-1)N(N-K)}{K} \approx \sigma^2(N^*),$$

d.h. beide Verfahren besitzen bezüglich desselben (erwarteten) Stichprobenaufwands eine Varianz vergleichbarer Größenordnung.

Die Varianz der Schätzung für N kann dadurch verringert werden, daß man so lange Tiere wiederfängt, bis man erstmalig n markierte Individuen gefunden hat. Bezeichnet man diese Zufallsgröße mit Y_n ($1 \leq n \leq K$), so gilt hierfür

$$E(Y_n) = n \frac{N+1}{K+1}.$$

Eine elementare Überlegung zeigt nämlich, daß

$$E(Y_2 | Y_1 = k) = \frac{N+1-k}{K} + k \quad (1 \leq k \leq N-K+1)$$

gilt, da die bedingte Verteilung von $Y_2 - Y_1$ unter $Y_1 = k$ mit der Verteilung von Y_1 übereinstimmen muß, wenn man N durch $N-k$ und K durch $K-1$ ersetzt: im ersten "Durchgang" werden ja gerade k Tiere gefangen, von denen genau eins markiert ist, d.h. es bleiben für den zweiten "Durchgang" $N-k$ Tiere übrig, von denen noch $K-1$ markiert sind. Hieraus folgt

$$\begin{aligned} E(Y_2) &= \sum_{k=1}^{N-K+1} E(Y_2 | Y_1 = k) \cdot P(Y_1 = k) \\ &= \frac{N+1}{K} + (1 - 1/K)E(Y_1) = 2 \frac{N+1}{K+1}. \end{aligned}$$

Induktiv erhält man damit das gewünschte Ergebnis für $E(Y_n)$, wenn man mit der Rekursion

$$E(Y_i | Y_{i-1} = k) = \frac{N+1-k}{K+2-i} + k \quad (2 \leq i \leq n, i-1 \leq k \leq N-K+1)$$

arbeitet.

Setzt man jetzt analog

$$N_n^* = \frac{K+1}{n} Y_n - 1,$$

so erhält man also wieder einen erwartungstreuen Schätzer für N . Die Verteilung von Y_n kann auch in geschlossener Form angegeben werden; sie lautet

$$P(Y_n = k) = \frac{\binom{K}{n-1} \binom{N-K}{k-n}}{\binom{N}{k-1}} \cdot \frac{K-n+1}{N-k+1}, \quad k = n, n+1, \dots, n+N-K \quad (n \in \mathbb{N}).$$

Der linke Bruch entspricht dabei gerade der (hypergeometrischen) Wahrscheinlichkeit dafür, bei Entnahme von $k-1$ Tieren aus der Grundgesamtheit von N Tieren genau $n-1$ markierte zu finden; dies entspricht dem Verlauf des sequentiellen Verfahrens vor dem Auffinden des n -ten markierten Tieres. Der rechte Bruch beschreibt die Wahrscheinlichkeit dafür, bei Vorhandensein von dann noch $K-n+1$ markierten Tieren bei einem restlichen Bestand von $N-k+1$ ein solches zu finden.

Die Varianz $\sigma^2(N_n^*)$ kann im sequentiellen Fall explizit berechnet werden; es gilt hier

$$\sigma^2(N_n^*) = \frac{(K-n+1)(N+1)(N-K)}{n(K+2)}.$$

Im Beispiel der norwegischen Forellen erhält man für $n = 57$ gerade $\sigma^2(N_n^*) = 645.54$, also einen geringfügig größeren Wert als beim Petersen–Verfahren, allerdings mit dem Vorteil der Erwartungstreue des sequentiellen Verfahrens. Ersetzt man in der Approximation für die Varianz $\sigma^2(\hat{N})$ die Größe n (=Anzahl Tiere im Wiederfang) durch die Größe $n(N+1)/(K+1)$ (=erwartete Anzahl von gefangenen Tieren bis zum n -ten markierten Tier), so erhält man wieder

$$\sigma^2(\hat{N}) \approx \frac{(N-K)N(K-n)}{nK} \approx \frac{(K-n+1)(N+1)(N-K)}{n(K+2)} = \sigma^2(N_n^*),$$

d.h. auch in diesem Fall entsprechen sich die Größenordnungen der beiden Varianzen. Die Varianz $\sigma^2(N_n^*)$ kann auch kleiner als $\sigma^2(\hat{N})$ ausfallen: im Beispiel $N = 100$, $k = 20$, $n = 50$ (Petersen–Methode) gilt etwa $\sigma^2(N_n^*) = 404 < 433.4 = \sigma^2(\hat{N})$, wenn man im sequentiellen Verfahren $n = 10$ wählt.

Eine detailliertere Behandlung von Capture–Recapture Verfahren findet man z.B. in der Monographie von Seber (1982).

Für große Werte von N und K ist die Berechnung der bisher betrachteten Verteilungen relativ umständlich. Die hypergeometrische Verteilung kann in diesem Fall z.B. gut durch die Binomialverteilung approximiert werden. Bei der Schätzung von Populationsanteilen gilt etwa (siehe S. 6):

$$\begin{aligned} P(X = k) &= \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}} = \frac{K!(N-K)!n!(N-n)!}{k!(K-k)!(n-k)!(N-K-n+1)!N!} \\ &\approx \binom{n}{k} \frac{K^k (N-K)^{n-k}}{N^n} = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, \dots, n \end{aligned}$$

wenn $K \geq n$, wobei wieder $p = K/N$ den wahren Populationsanteil bezeichnet. Im Fall mehrerer Teilstichproben kann die gemeinsame Verteilung X_1, \dots, X_r gut durch eine entsprechende Produktverteilung aus stochastisch unabhängigen Binomialverteilungen approximiert werden (siehe S. 9ff.):

$$\begin{aligned} P(X_1 = k_1, \dots, X_r = k_r) &\approx \prod_{i=1}^r \binom{n_i}{k_i} p^{k_i} (1-p)^{n_i-k_i} \\ &= \left[\prod_{i=1}^r \binom{n_i}{k_i} \right] p^{\sum_{i=1}^r k_i} (1-p)^{n - \sum_{i=1}^r k_i} \end{aligned}$$

für $0 \leq k_i \leq n_i$, $1 \leq i \leq r$, wenn $n := \sum_{i=1}^r n_i \leq K$.

Bei den Capture–Recapture Experimenten geht die Verteilung von Y bzw. Y_n (sequentielles Verfahren) entsprechend in eine geometrische bzw. negative Binomialverteilung über:

$$P(Y = k) \approx p(1-p)^{k-1}, \quad P(Y_n = k) \approx \binom{k-1}{n-1} p^n (1-p)^{k-n}, \quad k = n, n+1, \dots$$

Ist die Größe $p = K/N$ klein im Verhältnis zu n mit $np = \lambda > 0$, so kann im Fall der Binomialverteilung eine weitere Approximation durch die Poisson–Verteilung sinnvoll sein; es gilt nämlich dann z.B. bei der Schätzung von Populationsanteilen

$$\begin{aligned} P(X = k) &\approx \binom{n}{k} p^k (1-p)^{n-k} = \frac{\lambda^k}{k!} \frac{n!}{n^k (n-k)!} \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &\approx e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots \end{aligned}$$

Für die negative Binomialverteilung ist eine Poisson–Approximation dagegen nur sinnvoll, wenn $1-p$ klein im Verhältnis zu n ist, also p nahe bei 1 liegt, so daß etwa mit $\lambda = n(1-p)$ folgt

$$\begin{aligned} P(Y_n = k) &\approx \binom{k-1}{n-1} p^n (1-p)^{k-n} = \frac{\lambda^{k-n}}{(k-n)!} \frac{(k-1)!}{(n-1)! n^{k-n}} \left(1 - \frac{\lambda}{n}\right)^n \\ &\approx e^{-\lambda} \frac{\lambda^{k-n}}{(k-n)!}, \quad k = n, n+1, \dots \end{aligned}$$

Mit Hilfe der Poisson-Approximation erhält man z.B. näherungsweise bei der Petersen-Methode

$$\begin{aligned} E\left(\frac{1}{X+1}\right) &\approx e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{(k+1)!} = \frac{e^{-\lambda}}{\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{k+1}}{(k+1)!} \\ &= \frac{e^{-\lambda}}{\lambda} (e^{\lambda} - 1) = \frac{1}{\lambda} (1 - e^{-\lambda}), \end{aligned}$$

also

$$E(\hat{N}) \approx (K+1) \frac{N}{K} \left[1 - \exp\left(-\frac{nK}{N}\right) \right] \approx N \left[1 - \exp\left(-\frac{(n+1)(K+1)}{N}\right) \right],$$

was die frühere Approximationsformel (S. 13) anschaulich erklärt.

Eine Poisson-Approximation kann auch bei dem Problem mehrerer Teilstichproben für die Schätzung von Populationsanteilen sinnvoll sein. Ist der Anteil $p = K/N$ nämlich klein im Verhältnis zu den n_i , so folgt wie zuvor ($\lambda_i = n_i p$):

$$\begin{aligned} P(X_1 = k_1, \dots, X_r = k_r) &\approx \prod_{i=1}^r \binom{n_i}{k_i} p^{k_i} (1-p)^{n_i - k_i} \\ &\approx \prod_{i=1}^r e^{-\lambda_i} \frac{\lambda_i^{k_i}}{k_i!}, \quad k_i = 0, 1, \dots, 1 \leq i \leq r. \end{aligned}$$

Sowohl für die Binomial- als auch für die Poisson-Approximation gelten die auf S. 8ff. angestellten Überlegungen zur Minimierung der Varianz entsprechend. Das Problem vereinfacht sich allerdings durch die (approximative) stochastische Unabhängigkeit der Komponenten ganz erheblich: in beiden Fällen hat die (approximative) Varianz-Kovarianzmatrix Σ für die Teilschätzer \hat{p}_i nämlich Diagonalf orm, d.h. es gilt

– unter Binomialverteilung:

$$\Sigma = p(1-p) \begin{pmatrix} \frac{1}{n_1} & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{n_2} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & \frac{1}{n_r} \end{pmatrix}$$

– unter Poisson-Verteilung:

$$\Sigma = p \begin{pmatrix} \frac{1}{n_1} & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{n_2} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & \frac{1}{n_r} \end{pmatrix} \quad (\lambda_i = n_i p)$$

mit derselben (approximativ) optimalen Gewichtung wie zuvor:

$$a_i = \frac{n_i}{n_1 + \dots + n_r}, \quad i = 1, \dots, r.$$

Die folgenden Tabellen enthalten einige Zahlenwerte für die Binomial- und Poisson-Approximation der hypergeometrischen Verteilung:

$N = 50, K = 20, n = 10$			
k	hypergeometrisch	Binomial	Poisson
0	.00292	.00604	.01831
1	.02785	.04031	.07326
2	.10825	.12093	.14652
3	.22592	.21499	.19536
4	.28005	.25082	.19536
5	.21508	.20065	.15629
6	.10340	.11147	.10419
7	.03063	.04246	.05954
8	.00533	.01061	.02977
9	.00049	.00157	.01323
10	.00001	.00010	.00529

$N = 50, K = 10, n = 5$			
k	hypergeometrisch	Binomial	Poisson
0	.31056	.32768	.36787
1	.43133	.40960	.36787
2	.20983	.20480	.18393
3	.04417	.05120	.06131
4	.00396	.00640	.01532
5	.00011	.00032	.00306

$N = 100, K = 30, n = 10$			
k	hypergeometrisch	Binomial	Poisson
0	.02291	.02824	.04978
1	.11270	.12106	.14936
2	.23723	.23347	.22404
3	.28116	.26682	.22404
4	.20757	.20012	.16803
5	.09963	.10291	.10081
6	.03145	.03675	.05040
7	.00643	.00900	.02160
8	.00081	.00144	.00810
9	.00005	.00013	.00270
10	.00000	.00000	.00081

$N = 100, K = 20, n = 5$			
k	hypergeometrisch	Binomial	Poisson
0	.31930	.32768	.36787
1	.42014	.40960	.36787
2	.20734	.20480	.18393
3	.04784	.05120	.06131
4	.00514	.00640	.01532
5	.00020	.00032	.00306

$N = 200, K = 20, n = 5$			
k	hypergeometrisch	Binomial	Poisson
0	.58716	.59048	.60653
1	.33361	.32804	.30326
2	.07162	.07289	.07581
3	.00724	.00810	.01263
4	.00034	.00045	.00157
5	.00000	.00001	.00015

$N = 500, K = 20, n = 5$			
k	hypergeometrisch	Binomial	Poisson
0	.81468	.81537	.81873
1	.17115	.16986	.16374
2	.01363	.01415	.01637
3	.00051	.00058	.00109
4	.00000	.00001	.00005
5	.00000	.00000	.00000

Die angegebenen Beispiele zeigen deutlich, daß die Approximationen sowohl durch die Binomial- als auch durch die Poisson-Verteilung im Zentralbereich der hypergeometrischen Verteilung (d.h. um den Erwartungswert herum) bezüglich des relativen Fehlers günstiger sind als in den Randbereichen, und daß die Güte der Approximation besser wird, wenn $n \ll K \ll N$ gilt (man vgl. etwa die Situationen $N = 50, K = 10, n = 5$ mit $N = 100, K = 20, n = 5$ und $N = 500, K = 20, n = 5$).

Transekt-Verfahren

Häufig sind bei Freilanduntersuchungen flächenbezogene Informationen von Interesse, z.B. die Anzahl von Brutvogelpaaren auf einer Insel oder die Dichte einer Population je km^2 . Das gesamte Untersuchungsgebiet kann dabei allerdings in der Regel nicht vollständig beprobt werden (z.B. aufgrund von Auflagen aus dem Naturschutz oder wegen der Gesamtgröße der Fläche). Man erhebt daher häufig behelfsweise Daten aus Begehungen längs einer oder mehrerer Linien (Transekten) im Gelände, von denen aus man entweder Beobachtungen links und rechts der Strecke durchführt, oder entlang derer man Auszählungen in ausgewählten Teilflächen kleinerer Größe vornimmt. Aufgrund der ermittelten Daten erfolgt dann i.a. eine Hochrechnung auf den insgesamt vorhandenen Bestand im Untersuchungsgebiet bzw. auf die zugrundegelegte Referenzflächengröße.

Wir wollen zunächst untersuchen, welche räumlichen Verteilungsmuster entstehen, wenn sich Individuen im Untersuchungsgebiet "rein zufällig" niederlassen. Wir betrachten dazu r räumlich nicht überlappende Parzellen A_1, \dots, A_r mit den Flächeninhalten $m(A_i)$, $i = 1, \dots, r$, die auch nebeneinander – z.B. in Quadrat- oder Rechteckform – angeordnet sein können. Die Gesamtfläche sei durch

$$A = \bigcup_{i=1}^r A_i \quad \text{mit} \quad m(A) = \sum_{i=1}^r m(A_i)$$

bezeichnet. Bei einer "rein zufälligen" räumlichen Aufteilung geschieht die Auswahl einer Parzelle A_i durch ein Individuum mit der Wahrscheinlichkeit

$$p_i = \frac{m(A_i)}{m(A)}, \quad i = 1, \dots, r,$$

d.h. im Verhältnis der Flächeninhalte: eine "kleine" Parzelle wird seltener ausgewählt als eine "große". Geht man zunächst von einer festen Populationsgröße N im Untersuchungsgebiet aus, so sind die Zufallsvariablen X_1, \dots, X_r , die die jeweiligen Anzahlen von Individuen in den Parzellen A_1, \dots, A_r angeben, binomialverteilt:

$$P(X_i = k) = \binom{N}{k} p_i^k (1 - p_i)^{N-k}, \quad k = 0, 1, \dots, N;$$

denn jede Parzelle kann mehrfach – von unterschiedlichen Individuen – besetzt werden. Die X_1, \dots, X_r sind aber nicht unabhängig voneinander, sondern negativ korreliert: ist eine Parzelle mit "vielen" Individuen besetzt, so "fehlen" dafür Individuen in den übrigen Parzellen. Die gemeinsame Verteilung der X_1, \dots, X_r ist dementsprechend eine *Multinomialverteilung*:

$$P(X_1 = k_1, \dots, X_r = k_r) = \binom{N}{k_1, \dots, k_{r-1}} p_1^{k_1} \dots p_r^{k_r}, \quad 0 \leq k_i \leq N, \quad \sum_{i=1}^r k_i = N,$$

wenn man davon ausgeht, daß sich alle N Individuen auf die r Parzellen verteilen. Dabei ist wieder (vgl. S. 9)

$$\binom{N}{k_1, \dots, k_{r-1}} = \frac{N!}{k_1! \cdots k_{r-1}! k_r!}$$

im betrachteten Fall $\sum_{i=1}^r k_i = N$. Für je zwei Zufallsvariablen X_i, X_j mit $i \neq j$ gilt dabei ähnlich wie auf S. 10

$$\begin{aligned} E(X_i X_j) &= \sum_{\substack{0 \leq k_i, k_j \leq N \\ k_i + k_j \leq N}} k_i k_j P(X_i = k_i, X_j = k_j) \\ &= \sum_{\substack{1 \leq k_i, k_j \leq N \\ k_i + k_j \leq N}} k_i k_j \binom{N}{k_i, k_j} p_i^{k_i} p_j^{k_j} (1 - p_i - p_j)^{N - k_i - k_j} \\ &= N(N-1) p_i p_j \sum_{\substack{0 \leq k_i - 1, k_j - 1 \leq N - 2 \\ (k_i - 1) + (k_j - 1) \leq N - 2}} \binom{N-2}{k_i - 1, k_j - 1} p_i^{k_i - 1} p_j^{k_j - 1} \times \cdots \\ &\quad \cdots \times (1 - p_i - p_j)^{(N-2) - (k_i - 1) - (k_j - 1)} \\ &= N(N-1) p_i p_j, \end{aligned}$$

woraus

$$\text{Kov}(X_i, X_j) = N(N-1) p_i p_j - N^2 p_i p_j = -N p_i p_j$$

folgt. Die Varianz-Kovarianzmatrix Σ von (X_1, \dots, X_r) lautet demnach hier

$$\Sigma = N \begin{pmatrix} p_1(1-p_1) & -p_1 p_2 & -p_1 p_3 & \cdots & -p_1 p_r \\ -p_2 p_1 & p_2(1-p_2) & -p_2 p_3 & \cdots & -p_2 p_r \\ \vdots & \vdots & \vdots & & \vdots \\ -p_r p_1 & -p_r p_2 & -p_r p_3 & \cdots & p_r(1-p_r) \end{pmatrix}.$$

Beim Begehen einer Transekte werden in der Regel nicht alle r Parzellen des Untersuchungsgebiets erfaßt werden können, sondern es wird i.a. nur eine Teilauswahl $X_{r_1}, \dots, X_{r_\ell}$ mit $\ell < r$ untersucht. Will man auf der Basis der Zählungen $X_{r_1}, \dots, X_{r_\ell}$ einen erwartungstreuen Schätzer \hat{N} für die unbekannte Populationsgröße N der Form

$$\hat{N} = \sum_{j=1}^{\ell} a_j X_{r_j}$$

konstruieren, d.h.

$$N = E(\hat{N}) = \sum_{j=1}^{\ell} a_j E(X_{r_j}) = N \sum_{j=1}^{\ell} a_j p_{r_j},$$

also

$$\sum_{j=1}^{\ell} a_j p_{r_j} = 1 \quad \text{bzw.} \quad \mathbf{a}^{tr} \mathbf{p} = 1,$$

wobei entsprechend $\mathbf{a} = (a_{r_1}, \dots, a_{r_\ell})^{tr}$ und $\mathbf{p} = (p_{r_1}, \dots, p_{r_\ell})^{tr}$ zu setzen sind, so ist wieder ein Lagrange-Problem zu lösen:

$$\min! \quad \sigma^2(\hat{N}) = \mathbf{a}^{tr} \Sigma^* \mathbf{a} \quad \text{unter} \quad \mathbf{a}^{tr} \mathbf{p} = 1,$$

(vgl. S. 9), was sich analog auf die Lösung des Gleichungssystems

$$2\Sigma^* \mathbf{a} - \lambda \mathbf{p} = \mathbf{0}$$

reduziert, woraus durch Einsetzen

$$\mathbf{a} = c \cdot \mathbf{1}$$

mit einer geeigneten Konstanten c folgt. Dabei bezeichnet Σ^* die Varianz-Kovarianzmatrix der Zufallsgrößen $X_{r_1}, \dots, X_{r_\ell}$. Als erwartungstreuer Schätzer mit minimaler Varianz ergibt sich somit

$$\hat{N} = \frac{m(A)}{m(B)} \sum_{j=1}^{\ell} X_{r_j}$$

mit

$$B = \bigcup_{j=1}^{\ell} A_{r_j}, \quad \text{also} \quad m(B) = \sum_{j=1}^{\ell} m(A_{r_j}).$$

Die Zufallsgröße $X_B := \sum_{j=1}^{\ell} X_{r_j}$ ist selbst wieder binomialverteilt mit Erwartungswert $E(X_B) = N \sum_{j=1}^{\ell} p_{r_j} = Nm(B)/m(A)$, so daß sich für die (minimale) Varianz des Schätzers ergibt:

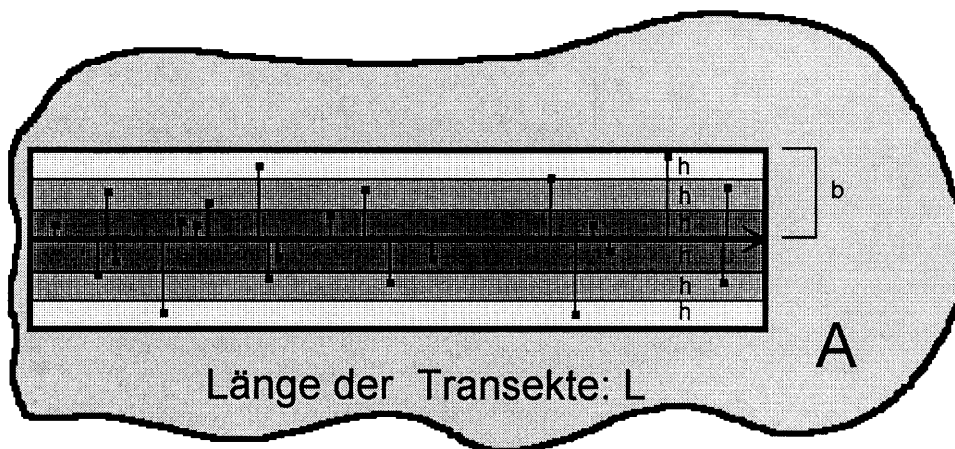
$$\begin{aligned} \sigma^2(\hat{N}) &= \left(\frac{m(A)}{m(B)} \right)^2 \sigma^2(X_B) = N \cdot \left(\frac{m(A)}{m(B)} \right)^2 \left[\frac{m(B)}{m(A)} \left(1 - \frac{m(B)}{m(A)} \right) \right] \\ &= N \left(\frac{m(A)}{m(B)} - 1 \right). \end{aligned}$$

Der Schätzer \hat{N} stellt also lediglich eine bereinigte Hochrechnung der Gesamtzahl $\sum_{j=1}^{\ell} X_{r_j}$ aller in den Parzellen $A_{r_1}, \dots, A_{r_\ell}$ gezählten Individuen dar. Bei einer Vollerhebung, d.h. $\ell = r$, reduziert sich die Varianz von \hat{N} nach obiger Rechnung natürlich auf Null.

Will man einen Schätzer für die Dichte D der Population, d.h. $D = N/m(A)$, so erhält man analog

$$\hat{D} = \frac{X_B}{m(B)} \quad \text{mit} \quad \sigma^2(\hat{D}) = D \left(\frac{m(A)}{m(B)} - 1 \right).$$

Ein häufig auftretendes Problem beim Begehen von Transekten besteht darin, daß die Zuverlässigkeit der Beobachtungen links und rechts der Transekte mit dem Abstand der Individuen vom Weg abnimmt. In der Regel werden dabei weiter entfernte Individuen leichter übersehen: die Gesamtpopulation wird daher systematisch unterschätzt. Eine Korrektur dieses Fehlers kann jedoch relativ einfach modellhaft berücksichtigt werden. Wir betrachten dazu den Fall, daß die Transekte die Länge L besitzt und samt einem $2b$ Längeneinheiten breiten, zur Transekte symmetrisch verlaufenden Streifen ganz innerhalb des Untersuchungsgebiets liegt.



Die Größe b kann beispielsweise durch die natürliche Sichtweite oder andere externe Bedingungen vorgegeben sein. $p(x)$ bezeichne die Wahrscheinlichkeit dafür, daß der Beobachter ein Individuum in der rechtwinkligen Entfernung x von der Transekte ($0 \leq x \leq b$) erkennt. Sinnvollerweise ist dabei $p(0) = 1$ anzunehmen, d.h. Individuen auf der Transekte werden mit Sicherheit erkannt. $p(x)$ wird in der Praxis in x monoton fallend sein. In erster Näherung stellen wir uns vor, daß der linke und rechte Streifen je in n Teilstreifen der Breite $h = b/n$ zerlegt sind, und in dem i -ten Streifen – von der Transekte aus gesehen – mit Wahrscheinlichkeit $p(ih)$ eventuell vorhandene Individuen übersehen werden. Bezeichnet $X_L(i)$ die Anzahl der im Streifen i links und $X_R(i)$ die Anzahl der im Streifen i rechts der Transekte vorhandenen Individuen, $Y_L(i)$ die Anzahl der im Streifen i links und $Y_R(i)$ die Anzahl der im Streifen i rechts der Transekte beobachteten Individuen, so sind $X_L(i)$, $Y_L(i)$ und $X_R(i)$, $Y_R(i)$ jeweils wieder binomialverteilt, und zwar mit den Erwartungswerten

$$E(X_L(i)) = E(X_R(i)) = h N \frac{L}{m(A)}$$

$$E(Y_L(i)) = E(Y_R(i)) = hp(ih) N \frac{L}{m(A)}.$$

Bezeichnen

$$X_L = \sum_{i=1}^n X_L(i), \quad X_R = \sum_{i=1}^n X_R(i),$$

$$Y_L = \sum_{i=1}^n Y_L(i), \quad Y_R = \sum_{i=1}^n Y_R(i)$$

die links und rechts der Transekte *vorhandenen* bzw. *beobachteten* Individuen, so folgt also

$$E(X_L) = E(X_R) = nhN \frac{L}{m(A)} = bN \frac{L}{m(A)}$$

$$E(Y_L) = E(Y_R) = \sum_{i=1}^n hp(ih)N \frac{L}{m(A)}.$$

Man beachte, daß im Gegensatz zu den Zufallsgrößen X_L und X_R die Zufallsgrößen Y_L und Y_R wegen der unterschiedlichen Parameter i.a. nicht mehr binomialverteilt sind. Läßt man jetzt die Streifenbreite h gegen Null streben, so konvergiert die Summe $\sum_{i=1}^n hp(ih)$ gegen das Integral $\int_0^b p(x) dx$; im Grenzwert ergibt sich somit

$$E(Y_L^*) = E(Y_R^*) = \left\{ \int_0^b p(x) dx \right\} N \frac{L}{m(A)},$$

wobei Y_L^* und Y_R^* die nach der kontinuierlichen Funktion $p(x)$ links und rechts der Transekte beobachteten Individuenzahlen bezeichnen. Faßt man nun noch die Größen Y_L^* und Y_R^* vermöge $Y^* = Y_L^* + Y_R^*$ zur Gesamtzahl der *beobachteten* Individuen im Transektbereich zusammen, so erhält man als erwartungstreuen Schätzer für N die Größe

$$\hat{N}^* = \frac{m(A)}{2L \int_0^b p(x) dx} Y^*.$$

Dieser Schätzer unterscheidet sich also von dem oben behandelten "klassischen" Schätzer lediglich durch den Faktor $\int_0^b p(x) dx$.

Die Funktion $p(x)$ muß in der Regel durch praktische Erfahrung oder Begleitexperimente ermittelt werden; sie kann als "Eichkurve" für den jeweiligen Beobachter angesehen werden. Im einfachsten Fall kann man $p(x)$ als linear annehmen, etwa

$$p(x) = 1 - \frac{x}{b}, \quad 0 \leq x \leq b;$$

in diesem Fall ist $\int_0^b p(x) dx$ gerade die Fläche des rechtwinkligen Dreiecks mit den Kantenlängen 1 und b , was den Wert $b/2$ ergibt. Es ist dann

$$\hat{N}^* = \frac{m(A)}{Lb} Y^*,$$

d.h. der Hochrechnungsfaktor ist doppelt so groß wie beim "klassischen" Schätzer. Weitere Informationen zu diesem Themenkreis findet man z.B. in dem Buch von Richter/Söndgerath (1990), S. 98ff. oder Mühlenberg (1993), S. 137ff.

Die bisherigen Betrachtungen gingen von einer konstanten Populationsgröße N aus. In der Natur ist aber die Populationsgröße in der Regel nicht konstant; so ändert sich beispielsweise der Brutvogelbestand einer Insel i.a. von Jahr zu Jahr. Solche Variabilitäten können z.B. in einem *Poisson-Modell* berücksichtigt werden; dabei wird die jeweilige Populationsgröße N als Zufallsgröße aufgefaßt, die einer Poisson-Verteilung mit Parameter (=Erwartungswert) $\lambda > 0$ genügt. Bezeichnet man auch in diesem Modell die in den einzelnen Parzellen A_1, \dots, A_r beobachteten Individuenzahlen wieder mit X_1, \dots, X_r , so sind diese sogar *stochastisch unabhängig* (und jeweils selbst wieder Poisson-verteilt): zunächst ist die gemeinsame *bedingte* Verteilung der X_1, \dots, X_r unter $N = n$ wie oben eine Multinomialverteilung, d.h.

$$P(X_1 = k_1, \dots, X_r = k_r \mid N = n) = \binom{n}{k_1, \dots, k_{r-1}} p_1^{k_1} \dots p_r^{k_r}$$

für $0 \leq k_i \leq n$, $\sum_{i=1}^r k_i = n$, so daß für die *unbedingte* gemeinsame Verteilung mit $\sum_{i=1}^r k_i = n$ folgt

$$\begin{aligned} P(X_1 = k_1, \dots, X_r = k_r) &= P(X_1 = k_1, \dots, X_r = k_r, N = n) \\ &= P(X_1 = k_1, \dots, X_r = k_r \mid N = n) \cdot P(N = n) \\ &= \binom{n}{k_1, \dots, k_{r-1}} p_1^{k_1} \dots p_r^{k_r} e^{-\lambda} \frac{\lambda^n}{n!} \\ &= \frac{p_1^{k_1} \dots p_r^{k_r}}{k_1! \dots k_r!} \exp\left(-\lambda \sum_{j=1}^r p_j\right) \lambda^{\sum_{j=1}^r k_j} \\ &= \frac{(p_1 \lambda)^{k_1} \dots (p_r \lambda)^{k_r}}{k_1! \dots k_r!} \prod_{j=1}^r e^{-p_j \lambda} \\ &= \prod_{j=1}^r \left[e^{-p_j \lambda} \frac{(p_j \lambda)^{k_j}}{k_j!} \right], \end{aligned}$$

woraus sowohl die Unabhängigkeit der X_j folgt (Produktstruktur der gemeinsamen Verteilung) als auch die Poisson-Verteilung mit Parameter $p_j \lambda$ für jedes X_j .

Diese auf den ersten Blick erstaunliche Tatsache erklärt, warum Poisson-Modelle als Basis-Modelle für "rein zufällige" Verteilungsmuster in der Ökologie besonders bedeutsam sind. Aus Sicht der Statistik vereinfachen sich die Rechnungen in solchen Modellen aufgrund der Unabhängigkeit beträchtlich: will man beispielsweise wieder eine erwartungstreue Schätzung minimaler Varianz $\hat{\lambda}$ für λ auf der Basis ausgewählter Anzahlen $X_{r_1}, \dots, X_{r_\ell}$ mit dem Ansatz

$$\hat{\lambda} = \sum_{i=1}^{\ell} a_{r_i} X_{r_i}$$

konstruieren, so hat man wegen der Unabhängigkeit lediglich die Lösung des einfacheren Problems

$$\min! \quad \sigma^2(\hat{\lambda}) = \mathbf{a}^{tr} \Sigma^* \mathbf{a} \quad \text{unter} \quad \mathbf{a}^{tr} \mathbf{p} = 1$$

mit der Diagonalmatrix

$$\Sigma^* = \lambda \begin{pmatrix} p_{r_1} & 0 & 0 & \dots & 0 \\ 0 & p_{r_2} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & p_{r_\ell} \end{pmatrix}$$

zu bestimmen, was direkt zu dem Schätzer

$$\hat{\lambda} = \frac{\sum_{i=j}^{\ell} X_{r_j}}{\sum_{j=1}^{\ell} p_{r_j}} = \frac{m(A)}{m(B)} \sum_{j=1}^{\ell} X_{r_j} = \frac{m(A)}{m(B)} X_B$$

führt, wenn wie oben wieder

$$\mathbf{a} = \begin{pmatrix} a_{r_1} \\ \vdots \\ a_{r_\ell} \end{pmatrix} \quad \mathbf{p} = \begin{pmatrix} p_{r_1} \\ \vdots \\ p_{r_\ell} \end{pmatrix} \quad B = \bigcup_{j=1}^{\ell} A_{r_j} \quad X_B = \sum_{j=1}^{\ell} X_{r_j}$$

gesetzt wird. Die Varianz des Schätzers ist in diesem Fall gegeben durch

$$\sigma^2(\hat{\lambda}) = \lambda \frac{m(A)}{m(B)}.$$

Die Form des Schätzers ist also identisch mit der Form des Schätzers für N bei fester Populationsgröße, besitzt aber demgegenüber eine etwas größere Varianz, bedingt durch die angenommene Variabilität des Populationsumfangs.

Das oben besprochene Transekten-Problem mit verminderter Zuverlässigkeit der Schätzung bei weiter entfernten Individuen läßt sich im Poisson-Modell völlig analog rechnen: auch hier ist die Zufallsgröße Y^* mit der Gesamtzahl der beobachteten Individuen im Transektbereich zu identifizieren; sie ist dabei selbst Poisson-verteilt mit Erwartungswert

$$E(Y^*) = 2 \left\{ \int_0^b p(x) dx \right\} \lambda \frac{L}{m(A)}$$

(vgl. S. 27), so daß analog zu \hat{N}^*

$$\hat{\lambda}^* = \frac{m(A)}{2L \int_0^b p(x) dx} Y^*$$

ein erwartungstreuer Schätzer für λ ist.

Abstands-Verfahren

Ein weiteres Verfahren zur Bestandsschätzung, das insbesondere in der Pflanzenökologie häufig eingesetzt wird, basiert auf den Abständen vom Beobachter zu den einzelnen Objekten im Untersuchungsgebiet, wie Bäumen, Büschen oder Sträuchern. Nehmen wir zunächst wieder an, daß eine feste Anzahl N von Holzgewächsen "rein zufällig" im Untersuchungsgebiet A verteilt ist, und wählen wir als Standort einen festen Punkt $\mathbf{x} \in A$, so läßt sich die Verteilungsfunktion F_Δ des Abstandes Δ von \mathbf{x} zum nächstgelegenen Holzgewächs durch die folgende Beziehung berechnen:

$$F_\Delta(r) = 1 - P(\Delta > r) \quad \text{mit} \quad P(\Delta > r) = \left(1 - \frac{m(K_r(\mathbf{x}) \cap A)}{m(A)}\right)^N, \quad r > 0,$$

wobei $K_r(\mathbf{x})$ den Kreis mit Mittelpunkt \mathbf{x} und Radius r bezeichnet. Das Ereignis $\{\Delta > r\}$ entspricht nämlich gerade dem Ereignis, daß das nächstgelegene Holzgewächs mehr als r Längeneinheiten vom Beobachter entfernt ist, d.h. keines der N im Gebiet A vorhandenen Holzgewächse des betrachteten Typs liegt innerhalb des Kreises $K_r(\mathbf{x})$; die angegebene Formel ergibt sich also aus der Binomialverteilung der Zufallsgröße "Anzahl der Gewächse in $K_r(\mathbf{x})$ " (vgl. S. 23). Für diejenigen Werte von r , für die der Kreis $K_r(\mathbf{x})$ ganz in A liegt, vereinfacht sich diese Beziehung wie folgt:

$$P(\Delta > r) = \left(1 - \frac{\pi r^2}{m(A)}\right)^N.$$

Für die übrigen Werte von r läßt sich der entsprechende Ausdruck i.a. nicht elementar vereinfachen, da hierfür der Flächeninhalt des Durchschnitts $K_r(\mathbf{x}) \cap A$ berechnet werden müßte. Die Verteilungsfunktion von Δ^2 ergibt sich hieraus zu

$$\begin{aligned} F_{\Delta^2}(z) &= 1 - P(\Delta^2 > z) = 1 - P(\Delta > \sqrt{z}) = 1 - \left(1 - \frac{m(K_{\sqrt{z}}(\mathbf{x}) \cap A)}{m(A)}\right)^N \\ &\approx 1 - \left(1 - \frac{\pi z}{m(A)}\right)^N, \quad 0 < z \leq \frac{m(A)}{\pi}, \end{aligned}$$

wobei die Darstellung exakt ist, solange z so klein bleibt, daß der Kreis $K_{\sqrt{z}}(\mathbf{x})$ ganz in A enthalten ist. (Dies trifft beispielsweise zu, wenn A selbst ein Kreis ist.) Näherungsweise erhält man aus dieser Darstellung für den Erwartungswert von Δ^2 :

$$\begin{aligned} E(\Delta^2) &= \int_0^\infty [1 - F_{\Delta^2}(z)] dz \approx \int_0^{m(A)/\pi} \left(1 - \frac{\pi z}{m(A)}\right)^N dz \\ &= -\frac{m(A)}{\pi} \frac{\left(1 - \frac{\pi z}{m(A)}\right)^{N+1}}{N+1} \Bigg|_0^{m(A)/\pi} = \frac{m(A)}{\pi(N+1)}, \end{aligned}$$

wobei die Approximation umso genauer ist, je größer $m(A)$, also die Fläche des Untersuchungsgebiets ist. Im Fall, daß A selbst ein Kreis ist, ist die Darstellung sogar exakt. Die Zufallsgröße

$$\hat{N} = \frac{m(A)}{\pi\Delta^2}$$

erscheint daher – zumindest in erster Näherung – als Schätzer für N plausibel. Bezeichnen wir mit

$$\lambda = \frac{N}{m(A)}$$

die Bewuchsdichte, so kann die Verteilungsfunktion von Δ^2 bei genügend großer Fläche $m(A)$ und konstanter Dichte λ auch wie folgt approximiert werden:

$$\begin{aligned} F_{\Delta^2}(z) &\approx 1 - \left(1 - \frac{\pi z}{m(A)}\right)^N = 1 - \left(1 - \frac{\pi z}{m(A)}\right)^{\lambda m(A)} \\ &\approx 1 - e^{-\pi\lambda z}, \quad z > 0; \end{aligned}$$

Die Verteilung von Δ^2 ist also approximativ eine Exponentialverteilung mit Parameter $\pi\lambda$ und Erwartungswert

$$E(\Delta^2) \approx \frac{1}{\pi\lambda} = \frac{m(A)}{\pi N},$$

was hinreichend gut mit dem vorigen Ergebnis übereinstimmt.

Diese Approximation kann im Rahmen eines allgemeineren Poisson-Modells exakt gemacht werden. Wir stellen uns dazu ein räumlich unbegrenztes Waldgebiet vor, das in unendlich viele nicht-überlappende Parzellen A_i , $i \in \mathbb{N}$, mit jeweils endlichem Flächeninhalt $0 < c \leq m(A_i) < \infty$ und endlicher Ausdehnung aufgeteilt ist. c ist hierbei eine vorgeschriebene Mindestgröße jeder Parzelle. In jeder Parzelle A_i befinden sich N_i Bäume, wobei N_i Poisson-verteilt ist mit dem Erwartungswert $\lambda \cdot m(A_i)$. Der Parameter $\lambda > 0$ entspricht dabei der Bewuchsdichte des Waldes (d.h. es ist $E(N_i) = \lambda$ für jede Parzelle A_i mit Flächeninhalt $m(A) = 1$). Die zufälligen Anzahlen N_i seien außerdem stochastisch unabhängig voneinander. In diesem Fall bleibt die Poisson-Eigenschaft auch für jede andere, beliebige Einteilung des Waldgebietes in unendlich viele nicht-überlappende Parzellen beschränkter Ausdehnung – sagen wir B_i , $i \in \mathbb{N}$ – erhalten, d.h. es gilt:

- Die Anzahlen M_i der in den Parzellen B_i befindlichen Bäume sind Poisson-verteilt mit Erwartungswert $E(M_i) = \lambda \cdot m(B_i)$;
- Die Anzahlen M_i der in den Parzellen B_i befindlichen Bäume sind stochastisch unabhängig.

Dies folgt unmittelbar aus den Ausführungen am Ende des vorigen Abschnitts (S. 28): Sei dazu B eine endliche Vereinigung von Parzellen B_{r_i} , $i = 1, \dots, \ell_1$ mit $\ell_1 \in \mathbb{N}$. Wegen der Überdeckungseigenschaft der ursprünglichen Parzellen A_i

findet man für B endlich viele Parzellen A_{s_j} , $j = 1, \dots, \ell_2$ mit $\ell_2 \in \mathbb{N}$, die B ganz überdecken, also mit

$$B \subseteq \bigcup_{j=1}^{\ell_2} A_{s_j} =: C.$$

Die Anzahl aller Bäume in der Menge C ist nun ebenfalls Poisson-verteilt mit Parameter $\lambda \cdot \sum_{j=1}^{\ell_2} m(A_{s_j})$. Die Anzahlen der Bäume in den neuen Parzellen B_{r_i} , $i = 1, \dots, \ell_1$ sind damit wieder bedingt multinomialverteilt und daher selbst stochastisch unabhängig und jeweils Poisson-verteilt mit den angegebenen Parametern. Bezeichnen wir nun mit $\xi(A)$ die Anzahl der Bäume in einem beliebigen Teilgebiet A des Waldes¹, so ist nach obigem die Verteilung von $\xi(A)$ unabhängig von der speziellen "Konstruktion" des Waldes mit

$$P(\xi(A) = k) = e^{-\lambda m(A)} \frac{[\lambda m(A)]^k}{k!}, \quad k = 0, 1, 2, \dots$$

und Erwartungswert

$$E(\xi(A)) = \lambda m(A).$$

Die Zufallsgröße $\xi(\cdot)$ wird in der mathematischen Literatur auch als *räumlicher Poisson-Prozeß* oder auch *Poisson-Wald* bezeichnet.

Wählen wir jetzt wieder einen Standort $\mathbf{x} \in A$ aus, so ist das Ereignis $\{\Delta > r\}$, also das Ereignis, daß der nächstgelegene Baum mehr als r Längeneinheiten vom Beobachter entfernt ist, wieder mit dem Ereignis, daß im Kreis $K_r(\mathbf{x})$ kein Baum vorhanden ist, identisch, so daß

$$P(\Delta > r) = P(\xi(K_r(\mathbf{x})) = 0) = e^{-\pi\lambda r^2}, \quad r > 0.$$

Dies entspricht dem oben erhaltenen approximativen Ergebnis für eine feste Baumzahl in einem beschränkten, aber flächenmäßig großen Waldstück.

Für die Momente von Δ erhalten wir in einem Poisson-Wald:²

$$\begin{aligned} E(\Delta) &= \frac{1}{2\sqrt{\lambda}}, & E(1/\Delta) &= \pi\sqrt{\lambda}, \\ E(\Delta^2) &= \frac{1}{\pi\lambda}, & E(1/\Delta^2) &= \infty. \end{aligned}$$

Der Beweis dieser Beziehungen beruht auf Integrationen mit der Dichte

$$f_{\Delta^2}(x) = \pi\lambda e^{-\pi\lambda x}, \quad x > 0$$

der Exponentialverteilung; z.B. gilt

$$E(\Delta) = \int_0^\infty \pi\lambda\sqrt{x}e^{-\pi\lambda x} dx = \frac{1}{\sqrt{\pi\lambda}}\Gamma\left(\frac{3}{2}\right) = \frac{1}{\sqrt{\pi\lambda}}\frac{\sqrt{\pi}}{2} = \frac{1}{2\sqrt{\lambda}},$$

¹ mathematisch präziser: einer "meßbaren" beschränkten Teilmenge $A \subset \mathbb{R}^2$

² $\Gamma(z) = \int_0^\infty x^{z-1}e^{-x} dx$ bezeichnet die Euler'sche Gamma-Funktion

$$E(1/\Delta) = \int_0^\infty \frac{\pi\lambda}{\sqrt{x}} e^{-\pi\lambda x} dx = \sqrt{\pi\lambda} \Gamma\left(\frac{1}{2}\right) = \pi\sqrt{\lambda},$$

$$E(1/\Delta^2) = \int_0^\infty \frac{\pi\lambda}{x} e^{-\pi\lambda x} dx = \infty.$$

Insbesondere das letzte Ergebnis zeigt, daß der dem Schätzer \hat{N} (S. 31) entsprechende Schätzer

$$\hat{\lambda} = \frac{\pi}{\Delta^2}$$

für die Bewuchsdichte *keinen* endlichen Erwartungswert besitzt und damit praktisch nicht zu gebrauchen ist. Man kann sogar zeigen, daß es auf der Basis von Δ allein *überhaupt keinen* erwartungstreuen Schätzer für λ geben kann: denn wäre etwa $g(\Delta)$ ein solcher Schätzer mit einer geeigneten Funktion g , so müßte

$$\lambda = E(g(\Delta)) = \pi\lambda \int_0^\infty g(x) e^{-\pi\lambda x} dx$$

gelten, also

$$1 = \pi \int_0^\infty g(x) e^{-\pi\lambda x} dx$$

sein für alle Werte von $\lambda > 0$, was nicht möglich ist. Man kann sich in diesem Fall damit behelfen, daß mehrere, voneinander unabhängige kürzeste Abstände vom Beobachter zu den Bäumen in disjunkten Teilgebieten des Waldes erhoben werden. So wird bei der PCQ-Methode (**P**oint-**C**entered-**Q**uarter-**M**ethod) ein fester Standort \mathbf{x} im Gelände ausgewählt und anschließend mit Hilfe eines Kompasses eine Einteilung des Gebiets in die durch die vier Himmelsrichtungen vorgegebenen "natürlichen" Quadranten vorgenommen. In jedem Quadranten Q_i , $i = 1, 2, 3, 4$, wird dann die Entfernung Δ_i von \mathbf{x} zum nächstgelegenen Baum ermittelt. In einem Poisson-Wald sind diese Abstände stochastisch unabhängig voneinander, und ihre Quadrate sind jeweils wieder exponentialverteilt:

$$\begin{aligned} P(\Delta_i > r) &= P(\xi(K_r(\mathbf{x} \cap Q_i)) = 0) \\ &= \exp(-\lambda m(K_r(\mathbf{x}) \cap Q_i)) = e^{-\frac{\pi\lambda}{4} r^2}, \quad r > 0, \quad i = 1, 2, 3, 4, \end{aligned}$$

d.h. Δ^2 ist exponentialverteilt mit Parameter $\pi\lambda/4$. Das in Mühlenberg (1993, S. 124) zitierte Verfahren benutzt dann als Schätzer für λ die Größe

$$\hat{\lambda} = \frac{1}{\bar{\Delta}^2} \quad \text{mit} \quad \bar{\Delta} = \frac{1}{4} \sum_{i=1}^4 \Delta_i \quad (\text{mittlere Distanz}),$$

woraus sich unmittelbar eine entsprechende Schätzung \hat{N} für die absolute Dichte von Bäumen in einem Waldstück A , also für $E(\xi(A)) = \lambda m(A)$ (vgl. S. 31), ergibt:

$$\hat{N} = \frac{m(A)}{\bar{\Delta}^2}.$$

Dieser Schätzer ist allerdings (für einen Poisson-Wald) nicht erwartungstreu; es gilt nämlich aufgrund der Jensen'schen Ungleichung für strikt konvexe Funktionen nach den obigen Momentenformeln:

$$E(\hat{\lambda}) = E\left(\frac{1}{\overline{\Delta^2}}\right) > \frac{1}{E(\overline{\Delta^2})} = 4\sqrt{\lambda/4} = \lambda,$$

d.h. dieser Schätzer hat einen positiven Bias: er überschätzt den Baumbestand systematisch. Eine erwartungstreue Alternative erhält man durch

$$\hat{\lambda}^* = \frac{3}{\pi \overline{\Delta^2}} \quad \text{mit} \quad \overline{\Delta^2} = \frac{1}{4} \sum_{i=1}^4 \Delta_i^2 \quad (\text{mittlere quadratische Distanz}),$$

bzw.

$$\hat{N}^* = \frac{3m(A)}{\pi \overline{\Delta^2}}.$$

Dies ergibt sich aus der (bekannten) Tatsache, daß die Summe unabhängiger identisch exponentialverteilter Zufallsgrößen eine Erlang-Verteilung besitzt; im vorliegenden Fall besitzt die Summe $S = \sum_{i=1}^4 \Delta_i^2$ also die Verteilungsdichte

$$f_S(x) = \frac{(\pi\lambda/4)^4}{3!} x^3 e^{-\pi\lambda x/4}, \quad x > 0.$$

Damit folgt aber

$$\begin{aligned} E\left(\frac{1}{S}\right) &= \int_0^\infty \frac{1}{x} f_S(x) dx = \int_0^\infty \frac{(\pi\lambda/4)^4}{3!} x^2 e^{-\pi\lambda x/4} dx \\ &= \frac{\pi\lambda}{12} \int_0^\infty \frac{(\pi\lambda/4)^3}{2!} x^2 e^{-\pi\lambda x/4} dx = \frac{\pi\lambda}{12} \end{aligned}$$

und somit

$$E(\hat{\lambda}^*) = E\left(\frac{12}{\pi S}\right) = \lambda.$$

Auf dieselbe Weise erhält man

$$E\left(\frac{1}{S}\right)^2 = \frac{(\pi\lambda/4)^2}{6} \int_0^\infty \frac{(\pi\lambda/4)^2}{1!} x e^{-\pi\lambda x/4} dx = \frac{(\pi\lambda)^2}{96}$$

und damit

$$\sigma^2\left(\frac{1}{S}\right) = \frac{(\pi\lambda)^2}{288}$$

bzw.

$$\sigma^2(\hat{\lambda}^*) = \frac{\lambda^2}{2}, \quad \sigma^2(\hat{N}^*) = \frac{(\lambda m(A))^2}{2}.$$

Die Standardabweichungen dieser Schätzungen sind also relativ hoch; größenordnungsmäßig liegen sie bei etwa 70% des zu schätzenden Parameters (λ bzw. $\lambda m(A)$). Entsprechendes gilt für den in Mühlenberg (1993) vorgeschlagenen Schätzer (vgl. hierzu auch Bonham (1989), S. 148ff.). Eine Verbesserung der Genauigkeit erhält man hier nur durch eine entsprechende Erhöhung der Anzahl der Beobachtungssektoren; teilt man etwa das Gebiet in $m > 4$ gleich große Sektoren ein, so ergeben sich entsprechende Schätzer:

$$\hat{\lambda} = \frac{m}{4\bar{\Delta}^2} \quad \text{mit} \quad \bar{\Delta} = \frac{1}{m} \sum_{i=1}^m \Delta_i \quad (\text{mittlere Distanz}),$$

$$\hat{\lambda}^* = \frac{m-1}{\pi\bar{\Delta}^2} \quad \text{mit} \quad \bar{\Delta}^2 = \frac{1}{m} \sum_{i=1}^m \Delta_i^2 \quad (\text{mittlere quadratische Distanz}).$$

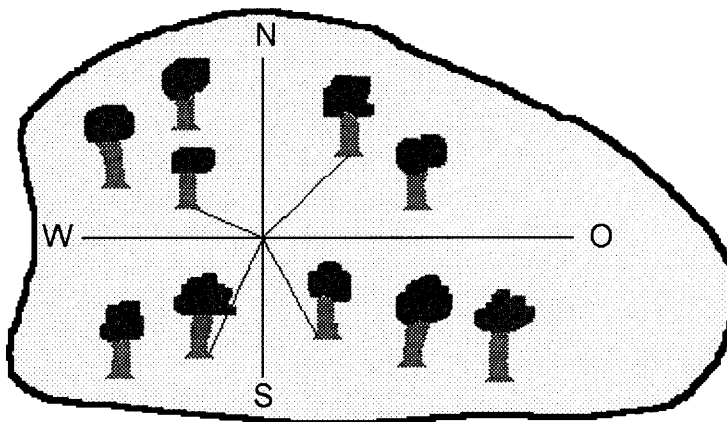
Hier ist ebenfalls nur der zweite Schätzer erwartungstreu; für ihn gilt analog

$$\sigma^2(\hat{\lambda}^*) = \frac{\lambda^2}{m-2}.$$

Die Summe $S = \sum_{i=1}^m \Delta_i^2$ besitzt in diesem Fall nämlich eine Erlang-Verteilung mit der Dichte

$$f_S(x) = \frac{(\pi\lambda)^m}{m^m (m-1)!} x^{m-1} e^{-\pi\lambda x/m}, \quad x > 0.$$

Im Gegensatz zum Poisson-Wald kann es bei Anwendung eines der Verfahren in einem realen Waldstück endlicher Ausdehnung eventuell (mit i.a. sehr geringer Wahrscheinlichkeit) passieren, daß in einem der Quadranten bzw. Sektoren keine Bäume vorhanden sind. In diesem Fall kann natürlich auch kein Minimalabstand ermittelt werden, so daß man entweder den Standort wechseln muß oder mit einer geringeren Zahl von Quadranten bzw. Sektoren arbeitet, was allerdings erneut zu einem (i.a. geringfügigen) Bias führt.



PCQ-Methode: Bestimmung kürzester Abstände

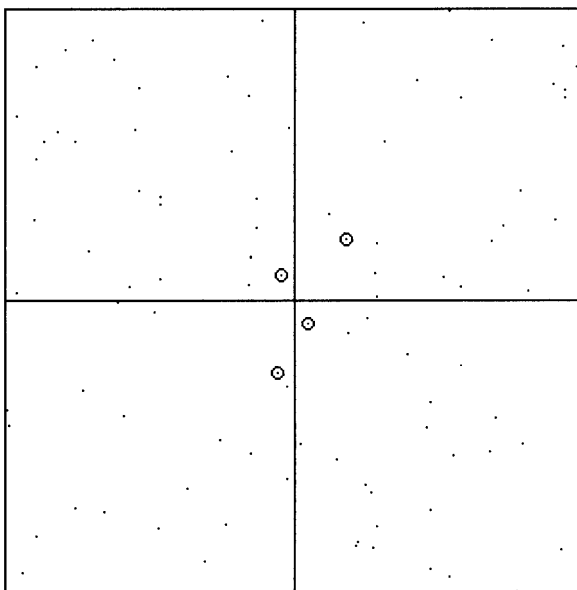
Die folgenden drei Graphiken zeigen das Ergebnis einer Simulationsstudie zur PCQ-Methode mit 100, 200 und 500 Durchläufen. Neben der Auswertung für das gegebene Bild sind jeweils auch der Mittelwert und die empirische Standardabweichung der Schätzer für die entsprechende Zahl von Durchläufen angegeben. Man sieht deutlich, daß der ursprüngliche PCQ-Schätzer $\hat{\lambda}$ zu hoch ausfällt, während der erwartungstreue modifizierte Schätzer $\hat{\lambda}^*$ schon recht nahe am "wahren" Wert liegt. Numerische Berechnungen zeigen, daß für den Erwartungswert und die Varianz von $\hat{\lambda}$ näherungsweise gilt:

$$E(\hat{\lambda}) \approx 1.268 \lambda, \quad \sigma^2(\hat{\lambda}) \approx 0.849 \lambda^2.$$

In der zweiten Zeile der Tabellen sind deshalb zum Vergleich noch jeweils der Mittelwert und die empirische Standardabweichung des korrigierten Schätzers $\hat{\lambda}/1.268$ angegeben.

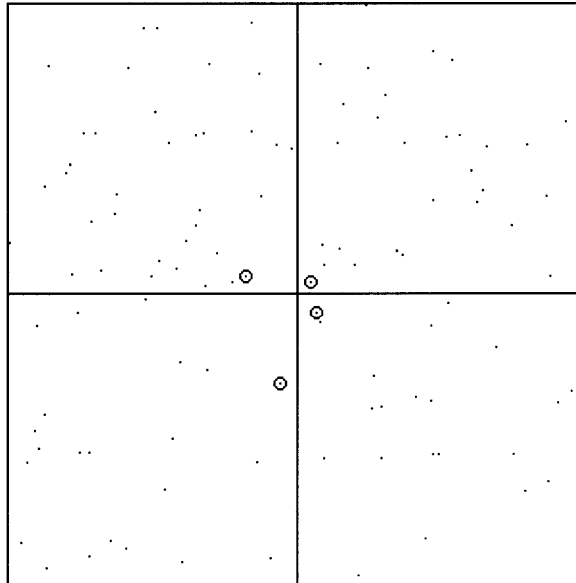
Die "Bäume" mit dem kürzesten Abstand zum "Beobachter" sind in jedem Quadranten eingekreist.

PCQ	122.36	Mittelwert:	112.86	Std.-Abw.:	77.25
PCQ (kor.)	96.50	Mittelwert:	89.00	Std.-Abw.:	60.97
PCQ (mod.)	95.33	Mittelwert:	86.71	Std.-Abw.:	49.63



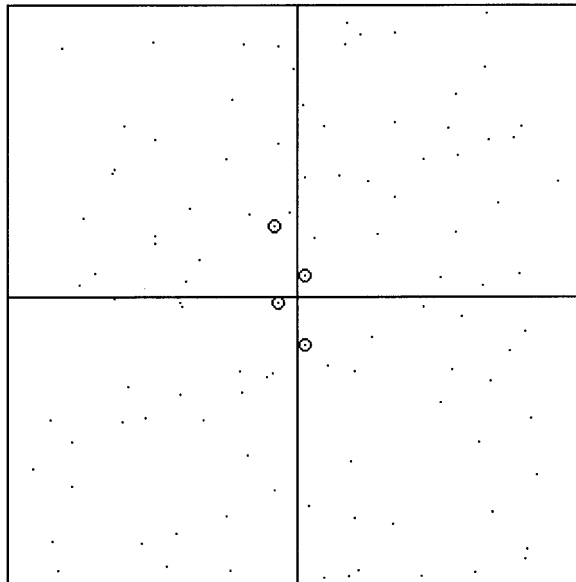
Nr. 100 Poisson-Modell Bewuchsdichte= 100 Baumzahl= 98

PCQ	146.80	Mittelwert:	125.44	Std.-Abw.:	74.38
PCQ (kor.)	115.77	Mittelwert:	98.93	Std.-Abw.:	58.70
PCQ (mod.)	102.18	Mittelwert:	99.20	Std.-Abw.:	55.53



Nr. 200 Poisson-Modell Bewuchsdichte= 100 Baumzahl= 111

PCQ	192.38	Mittelwert:	125.83	Std.-Abw.:	80.23
PCQ (kor.)	151.72	Mittelwert:	99.24	Std.-Abw.:	63.33
PCQ (mod.)	145.23	Mittelwert:	99.03	Std.-Abw.:	61.71



Nr. 500 Poisson-Modell Bewuchsdichte= 100 Baumzahl= 106

Versuchsplanung

Wenn man kein Ziel hat, bleibt man besser im Bett, es sein denn, der Weg ist das Ziel. Den meisten Forschern ist zu manchen Zeiten der Weg wichtiger als das Ziel. Sie sind neugierig und spielen herum, probieren dieses und jenes aus, einfach weil sie Spaß daran haben. Die Wissenschaft zehrt aber davon, daß sie gelegentlich fündig werden, das heißt, daß ihnen ein Problem klar wird und daß sie dann systematisch an der Lösung arbeiten – und an ein Ziel kommen. . . .

Eine präzise Antwort auf eine uninteressante Frage fördert die Karriere eines Wissenschaftlers genauso wenig wie keine Antwort auf eine interessante Frage. In der Praxis haben sich hingegen auch falsche Antworten auf interessante Fragen manchmal als nützlich für die Laufbahn des Autors erwiesen, besonders wenn es eine Weile dauerte, bis der Irrtum klar wurde. Voraussetzung ist allerdings, daß es sich nicht um bewußte Irreführung handelte, die, wenn entdeckt, einen Wissenschaftler für alle Zeiten für seinen Beruf untragbar macht. Falsche Antworten gehören zum normalen Gang der wissenschaftlichen Forschung, denn hier handelt man sich ständig von den schlechten zu den nächst besseren Hypothesen. Und wir dürfen froh sein, wenn wir soweit kommen, eine vorhandene Hypothese durch eine etwas bessere zu ersetzen. Auf den Anspruch, die absolute Wahrheit zu finden, verzichten wir besser von vorneherein. Dazu sind die Methoden der Naturwissenschaften nicht geeignet.

Was wir also in erster Linie suchen, sind "interessante" Fragen. Interessant ist meistens, was einen unmittelbaren Nutzen verspricht, was neu und daher überraschend ist, was Kontroversen erzeugt, was Kontroversen klärt oder was an das gängige Weltbild, an übergeordnete Theorien anknüpft.

Die Frage "Welche Stoffe wirken gegen AIDS?" ist lebenswichtig und wird deshalb auch in jeder Richtung intensiv erforscht. Der mögliche Nutzen einer Antwort liegt bei dieser Frage aus der Medizin genauso auf der Hand wie bei vielen Projekten im Bereich der Pflanzenzüchtung, Schädlingsbekämpfung oder **Ökologie**. . . .

Interessant wird eine Frage offenbar auch dadurch, daß sie einen Bezug zu einer übergeordneten Theorie oder einer allgemeinen Gesetzmäßigkeit hat, und diesen Bezug gilt es im Einzelfall zu finden.

Vielleicht gehen Sie von einer beobachteten Gesetzmäßigkeit aus und fragen, welche von den bekannten Theorien die passende Erklärung dafür abgeben könnte. Wenn zum Beispiel eine Tierart eine bestimmte Populationsdichte nie überschreitet, kommen als Erklärungen mindestens zwei Theorien in Betracht. Die eine macht den Raubfeinddruck dafür verantwortlich, die andere die Begrenztheit der Ressourcen wie Nahrung, Nistgelegenheiten und dergleichen. Die Frage ist: Welche Theorie trifft bei der untersuchten Population zu?

Theorien testet man, indem man aus ihnen Prognosen herleitet und dann nachsieht, ob sie zutreffen. Aus der Raubfeinddruck-Theorie würde man beispielsweise ableiten, daß die Beutetierpopulation anfangen wird zu wachsen, wenn man die Raubfeinde im Gebiet dezimiert. Das läßt sich nachprüfen. . . .

Meistens müssen wir Gesetzmäßigkeiten zuerst intuitiv erkennen, bevor wir sie ge-

zielt nachweisen und nach Erklärungen für sie suchen können. Die Fähigkeit zur Gestaltwahrnehmung spielt dabei eine wichtige Rolle. Sie setzt geduldiges, vorurteilsfreies Beobachten voraus. Gefragt ist die reine Neugier und nicht Selbstbestätigung, denn wer allzu sehr an eine Hypothese glaubt, übersieht leicht die Ausnahmen. Von der Notwendigkeit, das Ergebnis der Gestaltwahrnehmung mit **statistischen Methoden** zu überprüfen, war schon die Rede.

Über die Fähigkeit zur Gestaltwahrnehmung hinaus müssen wir in der Lage sein, Fragen in eine Form zu bringen, in der sie sich untersuchen lassen. Und außerdem sollten wir mögliche Einwände gegen unsere Schlußfolgerungen früh genug erkennen, um sie eventuell durch eine geeignetere Form der Datenerhebung (bessere Stichprobenauswahl, Experiment statt Korrelation, etc.) ausschließen zu können.

Wir können keine klare Antwort auf eine unklare Frage erwarten. . . . Unklare Fragen sind häufiger als man denkt. . . . Eine andere Art von Unklarheit kommt daher, daß die Frage zuviel Unbekanntes enthält. . . . Wieder andere Fragen sind in einem gewissen Sinn unklar, weil sie nach der Erklärung für eine Erscheinung fragen, ohne daß erkennbar ist, wie eine solche Antwort aussehen könnte. Oft ist das Problem einfach zu vielschichtig. Die Lösung läßt sich dann nicht in einem Satz formulieren.

[Leithilfen zur Planung von Experimenten:]

- Gewünschten Verallgemeinerungsbereich festlegen. Verfahren zur Erhebung einer "repräsentativen Stichprobe" ausarbeiten.
- Entscheiden, ob die Frage mit einem Experiment zu lösen ist, oder ob man sich mit quantitativen Beobachtungen ohne künstliche Eingriffe begnügen muß.
- Die Meßgrößen definieren und die Meßverfahren ausarbeiten (eventuell in Vorversuchen). Ebene der zu erhebenden Daten festlegen: Nominal-, Ordinal- oder Intervalldaten?
- :
- [Statistische Verfahren festlegen.]
- :
- Planen der Versuchs- oder Beobachtungsserien. Festlegen der Tage oder Tageszeiten der einzelnen Beobachtungen. Dauer einer Beobachtungssitzung oder eines einzelnen Experimentes. Reihenfolge, in der Experimente und Kontrollen durchgeführt werden sollen. Vorsorge gegen Abhängigkeit der Daten treffen (z.B. Pooling-Fehler vermeiden).
- :

Ich hoffe, daß aus dem bisher Besprochenen ganz klar geworden ist, daß es **unsinnig ist, mit dem Sammeln von Daten zu beginnen, bevor man weiß, wie man hinterher mit ihnen verfahren will**. Das unsystematische Sammeln von Notizen in Protokollbüchern ist sinnvoll in der Pilotphase eines Projektes, wo man die Tiere kennenlernen will und nach lohnenden Fragen Ausschau hält. Für den wissenschaftlich sauberen Nachweis von quantitativen Gesetzmäßigkeiten taugen solche Notizen in aller Regel nicht.

[aus Lamprecht (1992)]

Wir wollen uns in diesem Abschnitt hauptsächlich mit den *statistischen* Aspekten der Versuchsplanung auseinandersetzen. Die Ausführungen in den vorangegangenen Kapiteln haben gezeigt, daß die Anwendbarkeit aller in der Praxis verwendeten statistischen Verfahren auf bestimmten mathematischen Voraussetzungen beruht und diese Verfahren in der Regel mit Unsicherheiten behaftet sind, die z.B. durch Größen wie die Varianz quantifiziert werden können. Häufig existieren konkurrierende Methoden zur Untersuchung einer bestimmten Fragestellung; hier ist dann die Auswahl einer "optimalen" Methode gefragt, für die geeignete Kriterien heranzuziehen sind.

Nach der Klärung grundsätzlicher Fragen (vgl. obige Ausführungen) stellt sich meist die Frage nach dem notwendigen *Stichprobenumfang*. Im Zusammenhang mit der Schätzung von Populationsanteilen haben wir diesen Problemkreis bereits kurz diskutiert. Wesentliches Hilfsmittel dafür ist die *Tschebyscheff-Ungleichung*, die in allgemeiner Form so lautet:

Ist X eine Zufallsgröße mit Erwartungswert $E(X) = \mu$ und Varianz $\sigma^2(X) = \sigma^2$, so gilt für beliebige Zahlen $\varepsilon > 0$:

$$P(|X - \mu| > \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}.$$

Der Beweis dieser Ungleichung ist relativ einfach: betrachten wir dazu die Funktion

$$g(x) = \begin{cases} 0, & |x| \leq \varepsilon \\ \varepsilon, & |x| > \varepsilon, \end{cases}$$

so ist $|x| \geq g(x)$ für alle $x \in \mathbb{R}$ und somit

$$\sigma^2 = E(|X - \mu|^2) \geq E(g^2(X - \mu)) = \varepsilon^2 P(|X - \mu| > \varepsilon),$$

woraus die angegebene Ungleichung sofort folgt. Durch Umstellen der Terme ergibt sich auch die folgende, äquivalente Darstellung:

$$P(\mu \in [X - c\sigma, X + c\sigma]) \geq 1 - \frac{1}{c^2}$$

für alle Zahlen $c > 0$; hieraus wird noch einmal die besondere Bedeutung *erwartungstreuer* Schätzer ersichtlich: ist nämlich X ein solcher erwartungstreuer Schätzer für den Parameter μ , so stellt das (zufällige) Intervall $[X - c\sigma, X + c\sigma]$ gerade einen *Konfidenzbereich* zum (Unsicherheits-)Niveau $\alpha = 1/c^2$ dar. Je kleiner der Wert von c bzw. die Varianz σ^2 ausfällt, umso kleiner wird der zugehörige Konfidenzbereich und umso präziser damit auch die Schätzung von μ durch die Zufallsgröße X ; allerdings zu Lasten des Sicherheitsniveaus $1 - \alpha$, welches mit kleiner werdendem c ebenfalls abnimmt. Gibt man sich ein bestimmtes Sicherheitsniveau $1 - \alpha$ vor, so kann man die rechte Seite nach c auflösen und erhält

$$c = \frac{1}{\sqrt{\alpha}}.$$

Die Größe des Konfidenzbereichs hängt dann nur noch von der Varianz $\sigma^2 = \sigma^2(X)$ ab, die ihrerseits i.a. eine Funktion des Stichprobenumfangs n ist: $\sigma^2 = \sigma_n^2(X)$. Durch Auflösen nach n erhält man somit den notwendigen Stichprobenumfang, um bei einem gegebenen Sicherheitsniveau $1 - \alpha$ eine bestimmte Präzision der Schätzung (d.h. Größe des zugehörigen Konfidenzbereichs) zu erreichen. Die zugehörige Rechnung für die Schätzung von Populationsanteilen findet man z.B. auf S. 7f.

In vielen Fällen hängt die Varianz allerdings noch von dem unbekanntem Parameter μ selbst ab, etwa bei den Capture–Recapture Experimenten oder den Transekt– und Abstandsverfahren. In diesen Fällen kann man sich damit behelfen, eine Schätzung für die Varianz zu benutzen, wobei in der Regel das Sicherheitsniveau nicht mehr exakt, sondern nur noch approximativ eingehalten wird. Wir wollen diese Vorgehensweise kurz am Beispiel der sequentiellen Schätzung bei den Capture–Recapture Experimenten erläutern. Da der eigentliche Stichprobenumfang verfahrensbedingt zufällig ist, können wir nur eine möglichst gute Bestimmung der Zahl n , also der wiederzufangenden markierten Tiere, vornehmen. Nach den Ausführungen auf S. 17 ergibt sich bei genügend großer Gesamtpopulation N grob

$$\sigma^2(N_n^*) \approx \frac{(K - n)N(N - K)}{nK} \leq \frac{N^2}{n}.$$

Da die Anzahl N selbst unbekannt ist, ersetzen wir N durch die Schätzung N_n^* ; damit ist $(N_n^*)^2/n$ eine grobe approximative (zufällige!) obere Schranke für die Varianz. Das Intervall $N_n^*[1 - c/\sqrt{n}, 1 + c/\sqrt{n}]$ ist damit näherungsweise ein Konfidenzbereich zum Niveau α , mit $c = 1/\sqrt{\alpha}$. Für die hier betrachtete Fragestellung bietet sich die Vorgabe einer relativen Toleranz $\delta > 0$ an, d.h.

$$\frac{c}{\sqrt{n}} \leq \delta \quad \text{bzw.} \quad n \geq \frac{c^2}{\delta^2} = \frac{1}{\delta^2 \alpha}.$$

Gibt man beispielsweise eine relative Toleranz von $\delta = 0.1$ bei einem Sicherheitsniveau von $1 - \alpha = 0.9$ vor, so sollte $n \geq 1000$ sein; für $\delta = 0.2$ und $1 - \alpha = 0.8$ genügt dagegen bereits $n \geq 125$. Wegen $K \geq n$ hat man eine entsprechend große Zahl von Tieren im ersten Fang zu markieren. Ist die (unbekannte) Gesamtpopulation N klein, reichen natürlich kleinere Werte von n und K aus, allerdings sind diese Größen dann kaum im voraus planbar. Für das (nicht–sequentielle) Experiment von Dahl (S. 13ff) erhält man beispielsweise mit der dort angegebenen Schätzung für die Varianz den approximativen Konfidenzbereich

$$\left[337 - \frac{25}{\sqrt{\alpha}}, 337 + \frac{25}{\sqrt{\alpha}} \right].$$

Für $1 - \alpha = 0.9$ ergibt dies näherungsweise das Intervall $[258, 416]$, für $1 - \alpha = 0.95$ sogar $[225, 448]$.

Besonders kraß stellt sich die Situation bei der PCQ–Methode (S. 33ff) dar: hier liegt die Standardabweichung in der Größenordnung von etwa 70% des zu schätzenden Parameters, was bedeutet, daß für eine relative Toleranz von $\delta = 0.1$ bei einem

Sicherheitsniveau von $1 - \alpha = 0.9$ an mindestens $1/(2\alpha\delta^2) = 500$ verschiedenen Standorten Datenerhebungen vorzunehmen sind, wenn als Schätzung das arithmetische Mittel der Einzelschätzungen verwendet wird.

Eine Varianzreduktion kann auch durch unabhängige Versuchswiederholungen erreicht werden, wie es für die PCQ-Methode gerade angedeutet wurde. Nehmen wir wieder an, daß X ein erwartungstreuer Schätzer für den gesuchten Parameter $\mu = E(X)$ ist. Führen wir jetzt n voneinander unabhängige Experimente durch (parallel oder nacheinander), so erhalten wir die Einzelschätzer X_1, \dots, X_n , die wir folgendermaßen zu einem erwartungstreuen Schätzer \bar{X} zusammensetzen können:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{mit} \quad \sigma^2(\bar{X}) = \frac{1}{n^2} \sum_{i=1}^n \sigma^2(X_i) = \frac{\sigma^2}{n}.$$

Eine ähnliche Rechnung wie in den vorangegangenen Abschnitten zeigt hier, daß die *Gleichgewichtung* der einzelnen Schätzer tatsächlich auch zur minimalen Gesamtvarianz des zusammengesetzten Schätzers führt. Falls die Experimente voneinander *abhängig* sind (was in der Praxis eher der Fall ist) und man Informationen über die Kovarianzen oder Korrelationen der zugehörigen Einzelschätzer hat, so erhält man stattdessen den varianzminimalen Schätzer (vgl. S. 8ff.)

$$\bar{X} = \sum_{i=1}^n a_i X_i \quad \text{mit} \quad \mathbf{a} = (a_1, \dots, a_n)^{tr} = \frac{1}{\mathbf{1}\Sigma^{-1}\mathbf{1}} \Sigma^{-1} \mathbf{1},$$

wobei Σ wieder die Varianz-Kovarianzmatrix der Einzelschätzer bezeichnet.

Im (günstigen) Fall unabhängiger Experimente erhält man also mit der Tschebyscheff-Ungleichung den kleineren Konfidenzbereich

$$\left[\bar{X} - \frac{\sigma}{\sqrt{\alpha n}}, \bar{X} + \frac{\sigma}{\sqrt{\alpha n}} \right]$$

zur Sicherheitswahrscheinlichkeit $1 - \alpha$. Der Versuchsaufwand steigt demnach *quadratisch* mit der Präzision der Schätzung an, d.h. für eine durchschnittliche Halbierung des Konfidenzbereichs muß man z.B. viermal so viele Beobachtungen durchführen.

Eine Verkleinerung des Konfidenzbereichs ist auch möglich, wenn die Zufallsgröße \bar{X} (zumindest näherungsweise) normalverteilt ist. Aufgrund des zentralen Grenzwertsatzes ist das bei genügend großem n der Fall. Hierbei ergibt sich (ggf. näherungsweise) der i.a. kleinere Konfidenzbereich

$$\left[\bar{X} - c_\alpha \frac{\sigma}{\sqrt{n}}, \bar{X} + c_\alpha \frac{\sigma}{\sqrt{n}} \right]$$

zur Sicherheitswahrscheinlichkeit $1 - \alpha$, wobei

$$c_\alpha = \Phi^{-1} \left(1 - \frac{\alpha}{2} \right)$$

ist. $\Phi(z)$ bezeichnet hierbei die Verteilungsfunktion der Standard-Normalverteilung (d.h. mit Erwartungswert 0 und Varianz 1), die in den meisten Statistik-Lehrbüchern tabelliert ist. Die folgende Tabelle enthält einige Werte von c_α (zum Vergleich: $1/\sqrt{\alpha}$).

α	0.01	0.05	0.10	0.20	0.30
c_α	2.57	1.96	1.64	1.28	1.03
$1/\sqrt{\alpha}$	10.0	4.47	3.16	2.23	1.82

Ist die Varianz σ^2 nicht bekannt, so kann diese bei unabhängiger Stichprobenerhebung auch erwartungstreu geschätzt werden durch die *empirische Varianz*

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

Ersetzt man im entsprechenden Konfidenzbereich σ durch $\hat{\sigma}$, so sind die Größen c_α aus der sogenannten t -Verteilung zu bestimmen, wobei diese Größe noch vom Stichprobenumfang selbst abhängt, also eigentlich $c_\alpha = c_{n-1, \alpha}$ gilt. (Man spricht auch von der t -Verteilung mit $n-1$ *Freiheitsgraden*.) Die folgende Tabelle enthält einige Werte dieser Größen.

α	0.01	0.05	0.10	0.20
$c_{1, \alpha}$	63.6	12.7	6.31	3.07
$c_{2, \alpha}$	9.92	4.30	2.92	1.88
$c_{3, \alpha}$	5.84	3.18	2.35	1.63
$c_{5, \alpha}$	4.03	2.57	2.01	1.47
$c_{10, \alpha}$	3.16	2.22	1.81	1.37
$c_{20, \alpha}$	2.84	2.08	1.72	1.32

Für große Werte von n unterscheiden sich die Größen $c_{n-1,\alpha}$ und c_α aus der Normalverteilung erwartungsgemäß immer weniger.

Wendet man einmal die Tschebyscheff-Ungleichung und einmal die Normalapproximation auf die PCQ-Daten S. 36f. an, so erhält man als jeweilige (approximative) Konfidenzbereiche¹:

$n = 100$	Tschebyscheff		
α	0.01	0.05	0.10
PCQ (korr.)	[29.09 ; 177.63]	[70.15 ; 136.57]	[79.88 ; 126.84]
PCQ (mod.)	[34.76 ; 165.88]	[71.07 ; 129.63]	[79.59 ; 121.05]

$n = 200$	Tschebyscheff		
α	0.01	0.05	0.10
PCQ (korr.)	[50.54 ; 155.82]	[79.64 ; 126.72]	[86.54 ; 119.82]
PCQ (mod.)	[53.90 ; 149.86]	[80.73 ; 123.33]	[86.71 ; 117.05]

$n = 500$	Tschebyscheff		
α	0.01	0.05	0.10
PCQ (korr.)	[69.28 ; 132.96]	[86.88 ; 115.36]	[91.06 ; 111.18]
PCQ (mod.)	[70.75 ; 130.03]	[87.14 ; 113.64]	[91.02 ; 109.76]

¹ Die dort angegebenen empirischen Standardabweichungen wurden aus Konsistenzgründen im Programm mit dem Nenner n statt $n - 1$ ermittelt; diese Größen wurden für die hier durchgeführte Rechnung entsprechend korrigiert.

$n = 100$	Φ -Approx.		
α	0.01	0.05	0.10
PCQ (korr.)	[84.28 ; 122.44]	[88.81 ; 117.91]	[91.18 ; 115.54]
PCQ (mod.)	[83.48 ; 127.48]	[87.48 ; 113.16]	[89.57 ; 111.07]

$n = 200$	Φ -Approx.		
α	0.01	0.05	0.10
PCQ (korr.)	[89.66 ; 116.70]	[92.87 ; 113.49]	[94.55 ; 111.81]
PCQ (mod.)	[89.55 ; 114.21]	[92.48 ; 111.28]	[94.02 ; 109.74]

$n = 500$	Φ -Approx.		
α	0.01	0.05	0.10
PCQ (korr.)	[92.94 ; 109.30]	[94.88 ; 107.36]	[95.90 ; 106.34]
PCQ (mod.)	[92.78 ; 108.00]	[94.59 ; 106.19]	[95.53 ; 105.25]

Die approximative Verwendung der Werte c_α aus der Normalverteilung (statt der Werte $c_{n-1,\alpha}$ aus der t -Verteilung) ist hier aufgrund der großen Zahl n der Versuchswiederholungen gerechtfertigt¹.

¹ Die relative Abweichung beträgt in dem Rechenbeispiel maximal etwa 1%.

Konfidenzbereiche können auch zum Testen von Hypothesen über die Modellparameter verwendet werden. Will man beispielsweise testen, ob ein bestimmter Modellparameter $\mu = \mu_0$ vorliegt (einfacher Signifikanztest), so lautet die Entscheidungsregel:

- Annahme der Hypothese, falls $\mu_0 \in K_\alpha$
- Verwerfen der Hypothese, falls $\mu_0 \notin K_\alpha$.

Hierbei bezeichnet K_α einen geeigneten Konfidenzbereich zum Niveau α . Die Größe α entspricht dann gerade einer oberen Schranke für die *Fehlerwahrscheinlichkeit erster Art*, d.h. der Wahrscheinlichkeit dafür, daß man die Hypothese fälschlich verwirft, obwohl sie zutrifft.

Bei diesem einfachen Signifikanztest muß man in der Regel in Kauf nehmen, daß die *Fehlerwahrscheinlichkeit zweiter Art*, d.h. die fälschliche Annahme der Hypothese, obwohl sie nicht zutrifft, beliebig groß (bis hin zu 1) werden kann. Dies kann man – unter Beibehaltung der obigen Entscheidungsregel – i.a. nur vermeiden, wenn die Alternative: $\mu \neq \mu_0$ abgeschwächt wird, etwa in der Form: $\mu \notin [\mu_0 - \epsilon, \mu_0 + \epsilon]$ mit vorgegebenem $\epsilon > 0$. Die Fehlerwahrscheinlichkeit zweiter Art ist dann gegeben durch

$$P(\mu_0 \in K_\alpha) \quad \text{für} \quad |\mu - \mu_0| > \epsilon$$

(man beachte, daß das Intervall K_α von den Beobachtungen und damit indirekt auch von dem [wahren] Parameter μ abhängt!)

Für die PCQ-Methode ergibt sich damit bei n -facher Versuchswiederholung – zumindest näherungsweise für genügend "große" Werte von n – die Fehlerwahrscheinlichkeit zweiter Art zu

$$\begin{aligned} P(\mu_0 \in K_\alpha) &= P\left(\frac{\mu_0}{1 + c_\alpha/\sqrt{2n}} \leq \bar{X} \leq \frac{\mu_0}{1 - c_\alpha/\sqrt{2n}}\right) \\ &\approx \Phi\left(\left(\frac{\mu_0}{\mu} - 1\right)\sqrt{2n} + c_\alpha\right) - \Phi\left(\left(\frac{\mu_0}{\mu} - 1\right)\sqrt{2n} - c_\alpha\right) \\ &\approx 2c_\alpha \varphi\left(\left(\frac{\mu_0}{\mu} - 1\right)\sqrt{2n}\right) = \frac{2c_\alpha}{\sqrt{2\pi}} \exp\left(-\frac{n(\mu_0 - \mu)^2}{\mu^2}\right) \\ &\leq \frac{2c_\alpha}{\sqrt{2\pi}} \exp\left(-\frac{n\epsilon^2}{(\mu_0 - \epsilon)^2}\right) \end{aligned}$$

für $\mu \leq \mu_0 - \epsilon$ oder $\mu \geq \mu_0 + \epsilon$, wobei

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad x \in \mathbb{R},$$

die Dichte der Standard-Normalverteilung bezeichnet (*Gauß'sche Glockenkurve*). Hierbei wurde die in der Analysis gebräuchliche Approximation

$$\Phi(a+h) - \Phi(a-h) = \int_{a-h}^{a+h} \varphi(x) dx \approx 2h\varphi(a)$$

benutzt, die für "große" Werte von $|a|$ und "kleine" Werte von $h > 0$ sinnvoll ist. Wählt man z.B. $\alpha = 0.05$, $n = 100$, $\mu_0 = 100$ und $\epsilon = 10$, so ergibt der Test eine maximale Fehlerwahrscheinlichkeit zweiter Art von größenordnungsmäßig

$$\frac{2 \cdot 1.96}{\sqrt{2\pi}} \exp\left(-\frac{10^2 \cdot 100}{90^2}\right) = 0.455;$$

bei einem doppelt so großen Stichprobenumfang reduziert sich die maximale Fehlerwahrscheinlichkeit dagegen auf etwa

$$\frac{2 \cdot 1.96}{\sqrt{2\pi}} \exp\left(-\frac{10^2 \cdot 200}{90^2}\right) = 0.132.$$

Erst für 300 Versuchswiederholungen ergibt sich analog etwa

$$\frac{2 \cdot 1.96}{\sqrt{2\pi}} \exp\left(-\frac{10^2 \cdot 300}{90^2}\right) = 0.038,$$

also eine der Fehlerwahrscheinlichkeit erster Art größenordnungsmäßig entsprechende maximale Fehlerwahrscheinlichkeit zweiter Art.

In der angewandten Ökologie ist es häufig äußerst schwierig, wenn nicht unmöglich, Stichproben in den Größenordnungen zu erheben, die ein hohes Sicherheitsniveau beim Schätzen oder Testen von Modellparametern garantieren. Dies wird schon durch die obige Analyse der PCQ-Methode deutlich. Meist sind die Stichprobenumfänge auch durch äußere Einflüsse nach oben beschränkt, z.B. durch zeitaufwendige Laboruntersuchungen, die für die Versuchsauswertung benötigt werden, Witterungsverhältnisse oder auch Auflagen von Umweltschutzbehörden. Stichprobenumfänge können daher nicht immer im voraus optimal festgelegt werden. Man sollte aber in diesen Fällen unbedingt Konfidenzbereiche für die jeweils geschätzten Parameter bzw. Abschätzungen für die Fehlerwahrscheinlichkeit zweiter Art angeben, da sonst über die Aussagefähigkeit der Daten praktisch keine Informationen vorliegen. Dies ist vor allem für Umweltmonitoring-Programme bedeutungsvoll: jährliche Bestandsschwankungen z.B. von Brutvogelpaaren auf einer Insel (Zu- oder Abnahme), die auf geschätzten Populationsgrößen beruhen, müssen durchaus nicht immer den tatsächlichen Gegebenheiten entsprechen, sondern können sich leicht auch aufgrund der – leider unvermeidbaren – statistischen Fehler ergeben.

Eine gut lesbare Einführung in den Themenkreis des Schätzens und Testens von Parametern sowie in die allgemeinen Grundzüge der Versuchsplanung gibt neben dem Buch von Lamprecht (1992) auch die Monographie von Hartung, Elpelt und Klösener (1989).

2. Analyse und Vergleich von Lebensgemeinschaften

Physiologisch und verhaltenskundlich orientierte Ökologen beschäftigen sich hauptsächlich mit einzelnen *Organismen*. Koexistierende Individuen einer einzelnen Art besitzen jedoch Eigenschaften, wie Dichte, Geschlechtsverhältnis, Altersklassenstruktur, Raten von Geburt und Immigration sowie Mortalität und Emigration, die für *Populationen* charakteristisch sind. Wir erklären das Verhalten einer Population anhand des Verhaltens der Individuen, die sie enthält. Die Vorgänge auf Populationsniveau haben schließlich auch Folgen für das nächst höhere Niveau – das der *Lebensgemeinschaft*. Die Lebensgemeinschaft ist ein Verband von Populationen verschiedener Arten, die in Raum und Zeit zusammen vorkommen. Der Ökologe, der Lebensgemeinschaften untersucht, beschäftigt sich hauptsächlich mit den Ursachen der Verteilung natürlicher Artengemeinschaften, den Interaktionen zwischen Arten sowie physikalischen Einflüssen ihrer Umwelt, die diese Verteilung beeinflussen bzw. verursachen.

...

Ökologische Lebensgemeinschaften zu erforschen, das bedeutet oft, sich mit entmutigenden Problemen zu beschäftigen, weil die Datengrundlage sehr groß und komplex sein kann. Die Suche nach Mustern in der Struktur und Zusammensetzung der Lebensgemeinschaft stellt gewöhnlich einen ersten Schritt dar. Die Notwendigkeit, Methoden für die Beschreibung und den Vergleich von Lebensgemeinschaften auszuarbeiten, hat die Entwicklung der Ökologie von Lebensgemeinschaften beherrscht. Im Wesentlichen war dies eine Suche nach einfachen Wegen, um komplexe Systeme zu beschreiben.

...

Eine Möglichkeit, eine Lebensgemeinschaft zu charakterisieren besteht darin, einfach die vorhandenen Arten zu zählen oder aufzulisten. Dies hört sich nach einer einfachen Methode an, die uns ermöglicht, Lebensgemeinschaften durch ihren Artenreichtum zu beschreiben und zu vergleichen. In der Praxis ist dies aber oft überraschend schwierig, teilweise weil taxonomische Probleme auftreten, aber auch weil normalerweise nur eine **Stichprobe** der Organismen eines Gebietes gezählt werden kann. Die Zahl der registrierten Arten hängt somit von der Anzahl der genommenen Stichproben ab oder vom Volumen des Habitats, das untersucht wurde. Die häufigsten Arten sind wahrscheinlich in den ersten Probenahmen vertreten, und bei weiteren Stichproben werden seltenere Arten zur Liste hinzugefügt. An welchem Punkt hört man auf, weitere Stichproben zu nehmen? Idealerweise sollte der Untersuchende mit der Probenahme fortfahren, bis die Zahl der Arten ein Plateau erreicht. In der Praxis wird dies gewöhnlich nicht verwirklicht. Folglich kann der Artenreichtum unterschiedlicher Lebensgemeinschaften nur dann ordentlich verglichen werden, wenn er auf der gleichen Probengröße basiert (ausgedrückt als untersuchte Habitatfläche, für die Probenahme aufgewendete Zeit oder am besten als Anzahl der Individuen oder Module, die in der Stichprobe enthalten sind).

[aus Begon/Harper/Townsend (1991)]

Erfassung von Arten

Most ecological communities contain many species of organisms, and the species may vary greatly in their abundance from very common to very rare. Therefore, as soon as one attempts to study whole communities, rather than the interrelations among a few chosen species, the question immediately arises: how are the abundances of the different species distributed? If there are N individuals belonging to s species and the numbers of individuals in the respective species are N_1, N_2, \dots, N_s , have the N_j any consistent interrelationship, regardless of the type of community from which they come? Attempts to answer this question have led to the development of "species-abundance" curves. If it should turn out that one single form of **probability distribution** with a small number of parameters (say, two or three) fitted the data from the majority of observed communities, with only the parameter values varying from one community to another, interesting relationships might be discovered between the values of the parameters and the types of community they described.

...

In many collections it is found that singleton species (those represented by one individual) are numerous, often the most numerous. Species with successively more representatives, doubletons with 2, trebletons with 3, ..., and so on, are usually progressively less numerous. Roughly speaking, one often finds many rare species and a few abundant ones, although, of course, in terms of numbers of individuals those of the few common species far outnumber those of the many rare species. This frequently observed phenomenon has led to the method of tabulating **species-abundance data** customarily used: instead of listing the numbers of individuals in species 1, species 2, etc., we list the number of species, n_1 , represented by one member, ..., the number of species, n_r , represented by r members, ..., and so on. The n_r are, in fact, frequencies of frequencies.

[aus Pielou (1977)]

Wir wollen zunächst die in Begon, Harper und Townsend aufgeworfene Problematik der "Repräsentanz" von Arten einer Lebensgemeinschaft in einer Stichprobe vom Umfang n mathematisieren. Dazu nehmen wir an, daß die Lebensgemeinschaft aus s Spezies zusammengesetzt ist, von denen jede mit N_i Individuen, $i = 1, \dots, s$, vertreten ist. $N = \sum_{i=1}^s N_i$ ist also die Gesamtgröße der betrachteten Lebensgemeinschaft. Für $n \leq N$ sei die Zufallsgröße $Z_{i,n}$ definiert als

$$Z_{i,n} = \begin{cases} 1, & \text{falls Spezies } i \text{ in der Stichprobe vorhanden ist} \\ 0, & \text{sonst.} \end{cases}$$

Da es sich bei der Entnahme von Individuen aus der Lebensgemeinschaft um ein Experiment vom Typ *Ziehen ohne Zurücklegen* handelt, ergibt sich die Wahrschein-

lichkeit, in der Stichprobe kein Individuum der Spezies i vorzufinden, zu (vgl. S. 6)

$$P(Z_{i,n} = 0) = \frac{\binom{N - N_i}{n}}{\binom{N}{n}} = \frac{(N - N_i)(N - N_i - 1) \cdots (N - N_i - n + 1)}{N(N - 1) \cdots (N - n + 1)}.$$

Dabei gehen wir zunächst natürlicherweise davon aus, daß $N \geq 1$ ist, also die Lebensgemeinschaft überhaupt Individuen enthält.

Die Verteilung der Zufallsgröße

$$S_n = \sum_{i=1}^s Z_{i,n},$$

welche die insgesamt in der Stichprobe repräsentierten Arten angibt, läßt sich in der Regel allerdings nicht in einfacher Form explizit angeben. Für ihren Erwartungswert erhält man aber relativ leicht

$$\begin{aligned} f(n) &:= E(S_n) = \sum_{i=1}^s E(Z_{i,n}) = \sum_{i=1}^s (1 - P(Z_{i,n} = 0)) \\ &= \sum_{i=1}^s \left(1 - \prod_{j=1}^n \frac{N - N_i - j + 1}{N - j + 1} \right), \quad n = 1, \dots, N. \end{aligned}$$

Die durch f gegebene Funktion in Abhängigkeit von der Stichprobengröße heißt *Art-Repräsentanz-Kurve*. Sie gibt die mittlere Anzahl von Arten an, die in einer Stichprobe vom Umfang n repräsentiert sind, wenn die Auswahl der Individuen zufällig erfolgt. Sie besitzt die folgenden charakteristischen Eigenschaften:

$$f(1) = \sum_{i=1}^s \frac{N_i}{N} = 1 \quad \text{und}$$

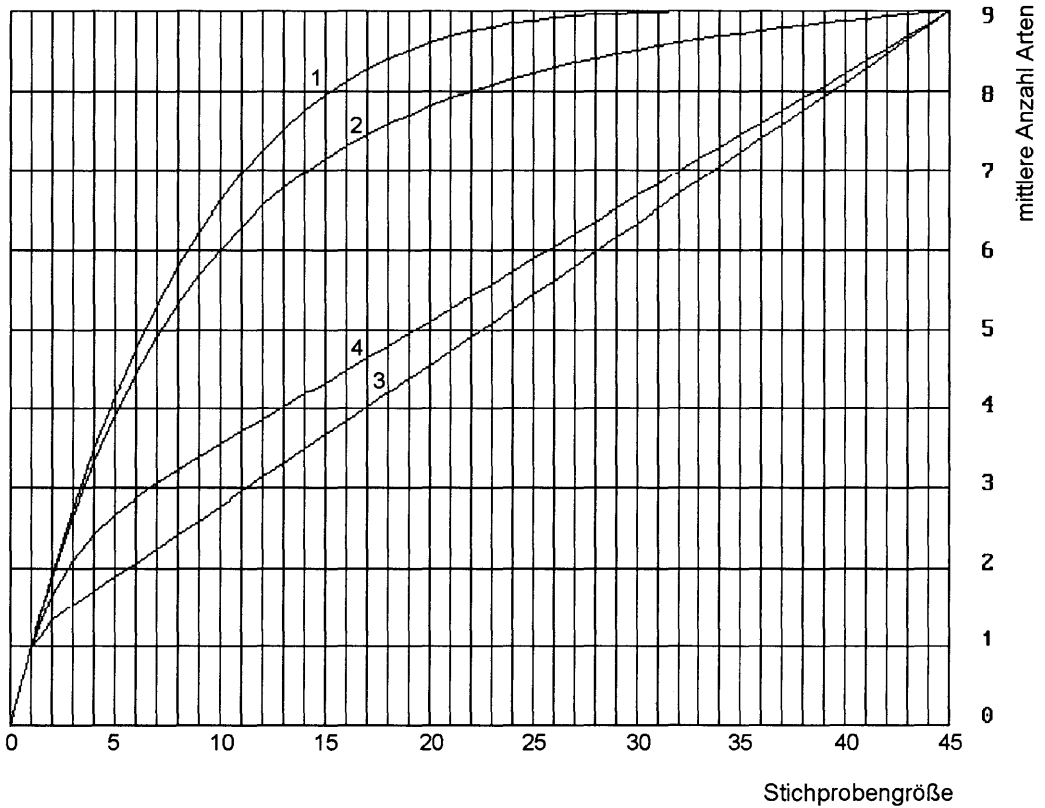
$$f(n) = s \quad \text{für } n \geq m := N + 1 - \min\{N_i \mid i = 1, \dots, s\}, \text{ falls alle } N_i \geq 1,$$

d.h. m ist der minimale Stichprobenumfang, bei dem mit Sicherheit alle Arten in der Stichprobe vertreten sind, wenn sie überhaupt in der Lebensgemeinschaft vorkommen.

Für die graphische Darstellung bietet es sich an, die aufeinanderfolgenden Punkte $(n - 1, f(n - 1))$ und $(n, f(n))$ für $n = 1, \dots, N$ miteinander zu verbinden, wobei natürlicherweise $f(0) := 0$ gesetzt wird. Der so modifizierte Graph von f ist dann monoton wachsend und konkav (im schwachen Sinn). Die folgenden Graphik zeigt vier verschiedene Art-Repräsentanz-Kurven für eine Lebensgemeinschaft aus neun Spezies mit 45 Individuen, und zwar für die Situationen

- 1) $N_i = 5, i = 1, \dots, 9$ (Gleichverteilung),
- 2) $N_i = i, i = 1, \dots, 9$ (schwache Ungleichverteilung),
- 3) $N_i = 1, i = 1, \dots, 8, N_9 = 37$ (Dominanz einer Art) und
- 4) $N_i = 1, i = 1, \dots, 7, N_8 = 19, N_9 = 19$ (Dominanz zweier Arten).

Die obere Linie entspricht dabei der Gleichverteilung; man beachte die Auswirkungen der unterschiedlichen Lebensgemeinschaftsstrukturen auf den jeweiligen Verlauf der Art-Repräsentanz-Kurven.

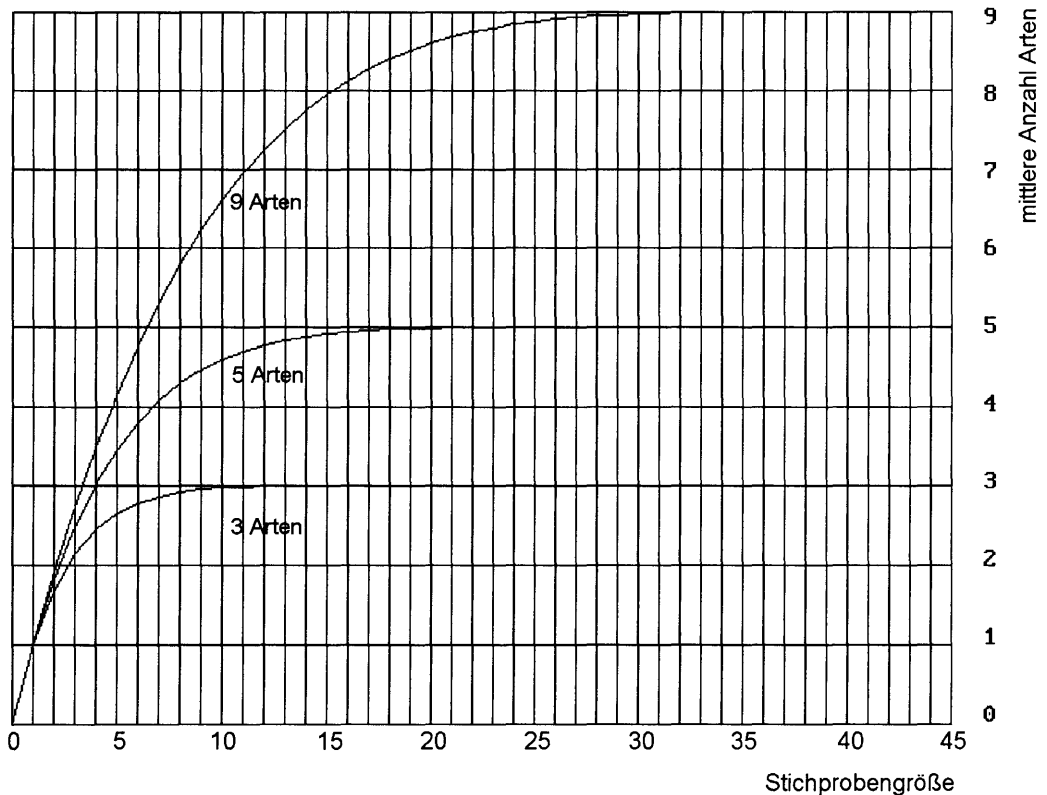


Art-Repräsentanz-Kurven für 9 Arten bei 45 Individuen

Die Graphik zeigt deutlich, daß zur Erfassung von 70% der Arten (im Mittel) ein Stichprobenumfang von 10 im Fall der Gleichverteilung ausreicht, während bei der schwachen Ungleichverteilung hierfür ein Stichprobenumfang von 12 erforderlich ist. Bei der Dominanz einer Art benötigt man dafür sogar einen Umfang von mindestens 31, bei der Dominanz zweier Arten (hier) 29.

Art-Repräsentanz-Kurven können natürlich auch in der Situation betrachtet werden, daß gewisse Arten in der Lebensgemeinschaft überhaupt nicht vorkommen, d.h. daß $N_i = 0$ gilt für gewisse i . In diesem Fall ist $f(n)$ genauso definiert; $N = \sum_{i=1}^s N_i \geq 1$ bezeichnet dabei weiterhin die Gesamtgröße der Lebensgemeinschaft. Der maximal erreichbare Wert von $f(n)$ beträgt dann allerdings nur noch $s^* = \#\{i \mid N_i > 0\} < s$, also die Gesamtzahl der *tatsächlich vorhandenen* Arten.

Die folgende Graphik zeigt 3 Art-Repräsentanz-Kurven für eine Gleichverteilung der Arten, wenn $s^* = 3, 5, 9$ beträgt.



Art-Repräsentanz-Kurven für 3, 5 und 9 Arten bei 45 Individuen

In der Praxis stellt sich häufig das Problem, daß die genaue Zahl der in einer Lebensgemeinschaft vorkommenden Arten nicht bekannt ist. Diese Anzahl muß dann geschätzt werden, damit man – wenigstens näherungsweise – einen Stichprobenumfang festlegen kann, bei dem ein vorgegebener Prozentsatz der Arten (im Mittel) auch tatsächlich erfaßt wird (vgl. hierzu die Graphik in Begon/Harper/Townsend (1991), S. 682). Da die Form der Kurven stark vom Grad der Homogenität der Artenverteilung in der Lebensgemeinschaft abhängt, kann man hier nur heuristisch vorgehen, d.h. man entscheidet sich für eine "einfache" parametrische Klasse von Funktionen, die monoton und konkav ist, und für die man anhand von Daten aus einer sequentiellen Stichprobennahme rechnerisch leicht eine Modellanpassung (z.B. mittels [nicht-linearer] Regression) vornehmen kann. Typische Klassen solcher Funktionen sind etwa

- Exponentialfunktionen: $f(x) = a(1 - e^{-bx})$, $a, b > 0$; $x \geq 0$ oder
- Hyperbeln: $f(x) = \frac{ax}{b+x} + cx$, $a, b > 0, c \geq 0$; $x \geq 0$.

Exponentialfunktionen bieten sich vor allem dann an, wenn ein hoher Grad an Homogenität in der Artenverteilung zu erwarten ist: bei einer Gleichverteilung der Arten und großer Individuenzahl sind diese Funktionen nämlich die "Grenzkurven" der Art-Repräsentanz-Kurven; denn bezeichnet $p_i = N_i/N$, $i = 1, \dots, s$ die Anteile der Art i in der Lebensgemeinschaft, so ist approximativ für große N (und

kleine n) wie oben

$$P(Z_{i,n} = 0) = \frac{(N - N_i)(N - N_i - 1) \cdots (N - N_i - n + 1)}{N(N - 1) \cdots (N - n + 1)} \approx (1 - p_i)^n$$

und demnach

$$E(S_n) \approx \sum_{i=1}^s (1 - (1 - p_i)^n) = \sum_{i=1}^s (1 - \exp(-n \lambda_i))$$

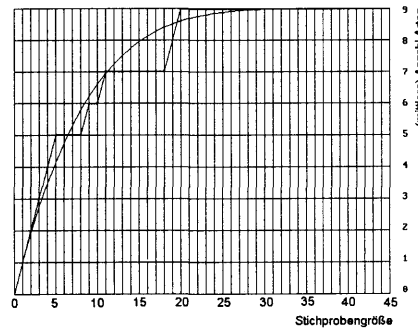
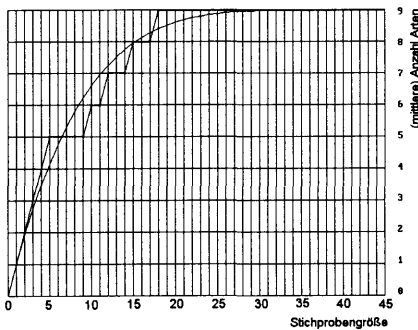
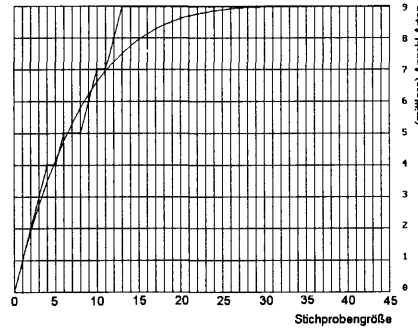
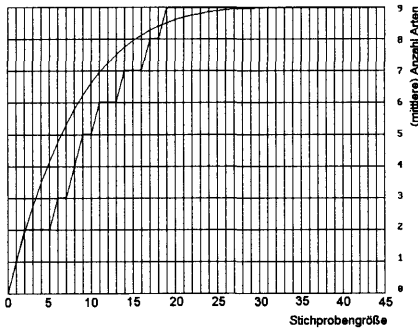
mit $\lambda_i = -\ln(1 - p_i) \geq 0$, $i = 1, \dots, s$. Bei einer Gleichverteilung der s^* in der Lebensgemeinschaft vorhandenen Arten bedeutet dies also

$$E(S_n) \approx a \cdot (1 - e^{-bn}) \quad \text{mit} \quad a = s^* \quad \text{und} \quad b = \lambda_{i_0} = -\ln(1 - p_{i_0}),$$

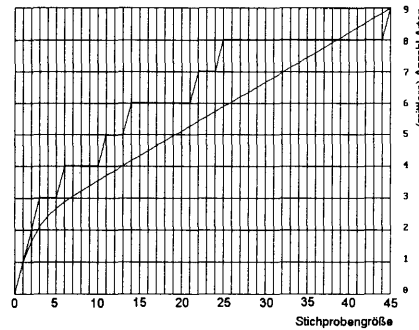
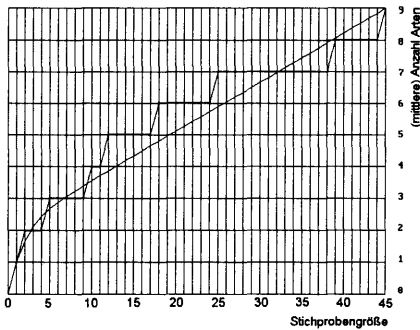
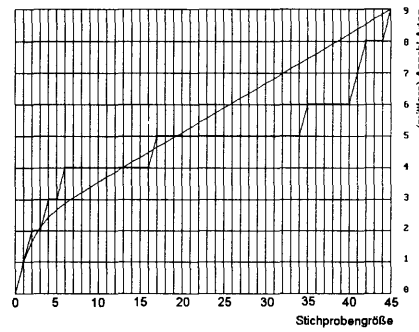
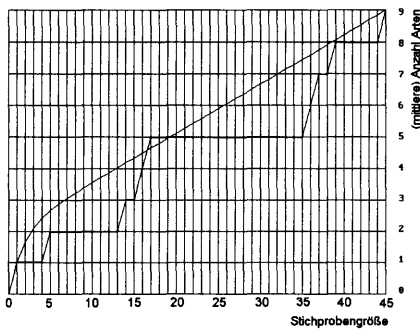
wobei i_0 ein Index ist mit $p_{i_0} > 0$. Arten i mit $\lambda_i = p_i = 0$ kommen in der Lebensgemeinschaft dabei nicht vor.

Man beachte, daß für beide Klassen von Funktionen der Parameter a jeweils ein "Schätzer" für die maximal vertretene Anzahl der Arten in der Lebensgemeinschaft, s^* , ist.

Die folgenden Graphiken zeigen jeweils vier verschiedene empirische Art-Repräsentanz-Kurven, und zwar einmal im Fall der Gleichverteilung der Arten sowie im Fall der Dominanz zweier Arten (vgl. S. 51) bei sequentieller Stichprobennahme. Man beachte, daß bei empirischen Art-Repräsentanz-Kurven die Konkavität i.a. verloren geht.



empirische Art-Repräsentanz-Kurven bei Gleichverteilung



empirische Art-Repräsentanz-Kurven bei Dominanz zweier Arten

Eine weitere mit der Art-Repräsentanz-Kurve in Zusammenhang stehende Funktion ist die *Art-Areal-Kurve*. Diese wird hauptsächlich dann betrachtet, wenn die (mittleren) Siedlungsdichten der einzelnen Arten je Flächeneinheit näherungsweise konstant sind und man an der Frage interessiert ist, wie groß das untersuchte Gebiet sein muß, damit ein gewisser Prozentsatz von Arten erfaßt wird (Bestimmung des sogenannten *Minimal-Areals*). Wir betrachten dieses Problem hier speziell für den Fall eines Poisson-Modells für jede einzelne Art, wie es früher (S. 31f.) schon einmal im Zusammenhang mit Abstandsverfahren betrachtet wurde. Dies bedeutet insbesondere, daß die entsprechenden Individuenzahlen der s möglichen Arten in der Lebensgemeinschaft für jede Fläche A mit der Größe $m(A)$ Poisson-verteilt sind, und zwar Art i mit dem Parameter $\lambda_i m(A)$, $i = 1, \dots, s$. λ_i bedeutet hier sinngemäß die Siedlungsdichte der Art i je Einheitsfläche. Man beachte, daß damit nicht vorausgesetzt wird, daß die Arten *unabhängig voneinander* vorkommen, sondern lediglich, daß die Abundanzen jeder einzelnen Art für sich genommen in nicht-überlappenden Teilgebieten unabhängige, Poisson-verteilte Zufallsgrößen sind. Dies ist eine natürliche Verallgemeinerung der gerade betrachteten Situation insofern, als bei gegebener Zahl N von Individuen in einem Gebiet A dann – im Mittel – $p_i \cdot N$ Individuen der Art i vertreten sind, wobei

$$p_i = \frac{\lambda_i}{\lambda_1 + \dots + \lambda_s}, \quad i = 1, \dots, s$$

gilt. Dies ergibt sich aus dem Zusammenhang zwischen der Multinomialverteilung und dem Poisson-Modell, wie er z.B. auf S. 28 untersucht wurde: bezeichnet etwa $X_i(A)$ die beobachtete Anzahl von Individuen der Art i im Areal A und $\mu_i := \lambda_i m(A)$, $i = 1, \dots, s$ sowie $X(A) = \sum_{i=1}^s X_i(A)$ den Gesamtumfang der Individuen in der Lebensgemeinschaft und $\mu = \sum_{i=1}^s \mu_i$, so ist

$$P(X_1(A) = N_1, \dots, X_s(A) = N_s \mid X = N) = \frac{P(X_1(A) = N_1, \dots, X_s(A) = N_s)}{P(X(A) = N)}$$

$$= \frac{\prod_{i=1}^s \frac{e^{-\mu_i} \mu_i^{N_i}}{N_i!}}{e^{-\mu} \frac{\mu^N}{N!}} = \binom{N}{N_1, \dots, N_{s-1}} \prod_{i=1}^s \left(\frac{\mu_i}{\mu}\right)^{N_i} = \binom{N}{N_1, \dots, N_{s-1}} \prod_{i=1}^s p_i^{N_i}$$

für $N = \sum_{i=1}^s N_i$, d.h. die bedingte gemeinsame Verteilung der Abundanzen der Arten bei gegebener Gesamtpopulation von $X(A) = N$ ist multinomialverteilt mit den angegebenen Parametern p_1, \dots, p_s . Insbesondere ist also

$$E(X_i(A) \mid X(A) = N) = p_i N, \quad i = 1, \dots, s$$

wie behauptet. Man beachte, daß die obigen Rechnungen nicht von dem Gebiet A oder dessen Größe $m(A)$ abhängen. Definiert man wieder analog

$$Z_i(A) = \begin{cases} 1, & \text{falls Spezies } i \text{ in } A \text{ vorhanden ist} \\ 0, & \text{sonst} \end{cases} \quad \text{sowie} \quad S(A) = \sum_{i=1}^s Z_i(A),$$

so ergibt sich die Art-Areal-Kurve entsprechend zu

$$g(A) := E(S(A)) = \sum_{i=1}^s E(S_i(A)) = \sum_{i=1}^s (1 - P(S(A) = 0))$$

$$= \sum_{i=1}^s (1 - \exp(-\lambda_i m(A))).$$

Die Art-Areal-Kurve wird allerdings meist als Funktion h der Flächengröße $x = m(A)$ statt als Funktion des Gebiets A selbst angegeben (vgl. Krebs (1985), S. 551ff. oder Mühlenberg (1993), Kapitel 8.5.6), d.h. es ist

$$h(x) = \sum_{i=1}^s (1 - e^{-\lambda_i x}) = \sum_{i=1}^{s^*} (1 - e^{-\lambda_i x}), \quad x > 0,$$

wegen $1 - e^{-\lambda_i x} = 0$ für Arten i mit $\lambda_i = 0$ (d.h. diese Arten kommen in der Lebensgemeinschaft nicht vor). In dieser Form gleicht die Art-Areal-Funktion h der oben betrachteten Art-Repräsentanzfunktion f : sie ist ebenfalls monoton und konkav mit s^* – der Anzahl der tatsächlich in der Lebensgemeinschaft vertretenen Arten – als asymptotisch größtem Wert. Analytisch ergeben sich diese Eigenschaften aus

$$h'(x) = \sum_{i=1}^s \lambda_i e^{-\lambda_i x} > 0 \quad (\text{Monotonie})$$

$$h''(x) = - \sum_{i=1}^s \lambda_i^2 e^{-\lambda_i x} < 0 \quad (\text{Konkavität})$$

für $x > 0$ sowie $\lim_{x \rightarrow \infty} h(x) = s^*$ wegen $\lim_{x \rightarrow \infty} e^{-\lambda_i x} = 0$ für alle Arten i mit $\lambda_i > 0$.

In praktischen Anwendungen werden häufig Flächengrößen im Bereich mehrerer Zehnerpotenzen erfaßt (vgl. Krebs (1985), S. 551ff.). Es ist daher sinnvoll, die x -Achse logarithmisch zu teilen. Wegen der großen Zahl von Arten in ausgedehnten Arealen (bis hin zu Ländergröße oder Kontinenten) wird i. a. auch die y -Achse logarithmisch geteilt. Mathematisch erhält man die entsprechenden Funktionsgraphen durch die Transformation

$$a^y = h(a^x), \quad x \in \mathbb{R}$$

mit Exponent $a > 1$ (typischerweise: $a = 10$) bzw. äquivalent

$$y = \log_a a^y = \log_a (h(a^x)), \quad x \in \mathbb{R},$$

wobei x und y sich auf eine linear geteilte Skala beziehen. Die Steigungen der Funktionsgraphen (in der linearen Skala) ergeben sich damit zu

$$y' = \frac{h'(a^x) a^x}{h(a^x)}.$$

Für das Poisson-Modell ergibt sich daraus der Ausdruck

$$y' = \frac{\sum_{i=1}^s \lambda_i a^x \exp(-\lambda_i a^x)}{\sum_{i=1}^s (1 - \exp(-\lambda_i a^x))} \approx 1 - \frac{a^x \sum_{i=1}^s \lambda_i^2}{2 \sum_{i=1}^s \lambda_i}$$

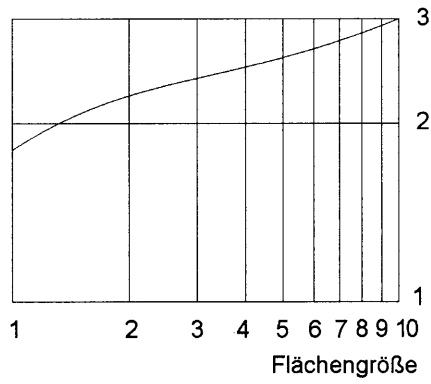
für "kleine" Werte von $a^x \max\{\lambda_i \mid i = 1, \dots, s\}$. Wählt man als Nullpunkt den Punkt (1,1) in der $\log - \log$ -Skala (d.h. den gewöhnlichen Nullpunkt in der linearen Skala), so ergibt sich dort als "Anfangssteigung" näherungsweise

$$y'(0) \approx 1 - \frac{1 \sum_{i=1}^s \lambda_i^2}{2 \sum_{i=1}^s \lambda_i} < 1,$$

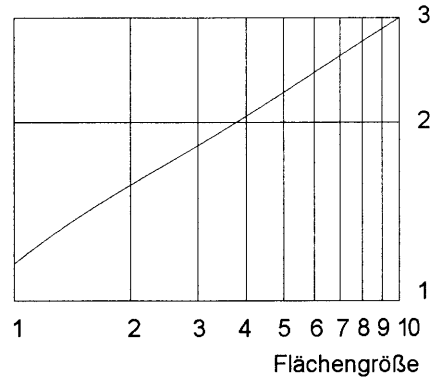
falls $\max\{\lambda_i \mid i = 1, \dots, s\}$ genügend klein ist.

Die folgenden sechs Diagramme geben die Art-Areal-Kurven für $s = 10$ Arten in der $\log_{10} - \log_{10}$ -Skala mit verschiedenen Wahlen der λ_i wieder.

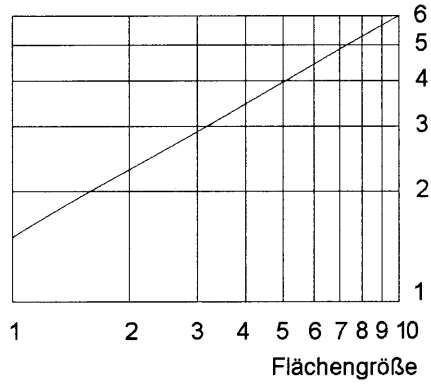
Kurve Nr.	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}
1	2.0	1.5	0.05	0.01	0.005	0.005	0.005	0.005	0.005	0.005
2	2.0	0.2	0.05	0.01	0.005	0.005	0.005	0.005	0.005	0.005
3	1.8	0.2	0.1	0.05	0.05	0.05	0.05	0.05	0.05	0.05
4	2.7	0.3	0.15	0.1	0.1	0.1	0.1	0.1	0.1	0.1
5	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
6	10.0	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05



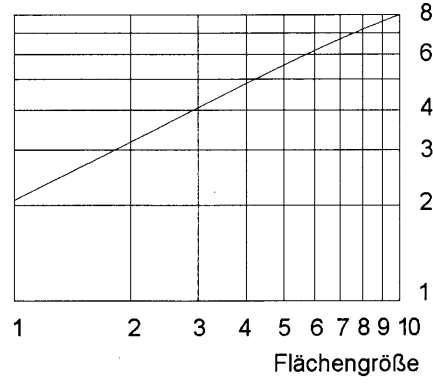
Kurve 1



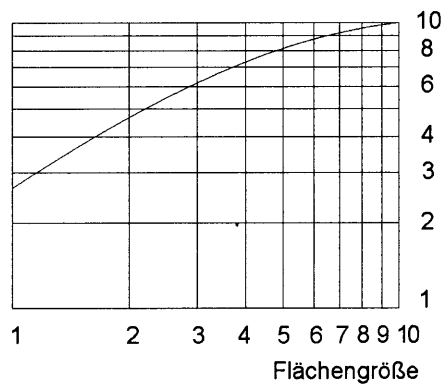
Kurve 2



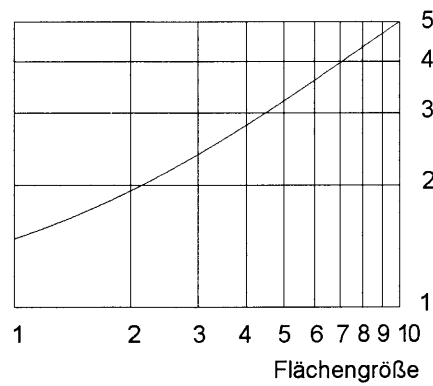
Kurve 3



Kurve 4



Kurve 5



Kurve 6

Man erkennt, daß im doppelt-logarithmischen Maßstab offensichtlich unterschiedliche Krümmungseigenschaften der Kurven auftreten können: so ergibt sich für Kurve 1 etwa ein s-förmiger Verlauf, für die Kurven 2, 3 und 4 (näherungsweise) ein linearer Verlauf, für die Kurve 5 ein konkaver und die Kurve 6 sogar ein konvexer Verlauf.

In der ökologischen Literatur wird vielfach davon ausgegangen, daß "reale", d.h. empirisch gewonnene Art-Areal-Kurven im doppelt-logarithmischen Maßstab einen – zumindest näherungsweise in bestimmten Bereichen – linearen Verlauf haben (vgl. hierzu Krebs (1985), S. 551 ff.). Allerdings werden auch Kurven beobachtet, die in ihrem Verlauf eher Kurve 1 ähneln (vgl. Krebs (1985), S. 553, Figure 24.9). Wir wollen hier kurz diskutieren, welche *mathematischen* Hintergründe solche Kurvenverläufe haben, wenn man – wie oben – ein Poisson-Modell unterstellt.

Für Kurve 1 ist typisch, daß zwei Arten dominant sind ($\lambda_1 = 2.0$, $\lambda_2 = 1.5$) und zugleich viele seltene Arten auftreten ($\lambda_i = 0.005$ für $i = 5, \dots, 10$). Kurve 2 zeigt denselben Effekt, allerdings erheblich schwächer ausgeprägt und erscheint eher linear in ihrem Verlauf. Die Dominanz weniger Arten führt dazu, daß diese bereits bei kleinen Flächengrößen relativ schnell vollständig erfaßt werden (starkes Ansteigen der Kurve am "Anfang"), während die vielen seltenen Arten – beinahe schlagartig – erst bei entsprechend großen Arealen erfaßt werden (starkes Ansteigen der Kurve am "Ende"). Man beachte, daß der optisch doch recht unterschiedliche Verlauf der Kurven 1 und 2 bereits durch eine Abundanzänderung in nur einer einzigen der zehn vertretenen Arten verursacht wird.

Die Kurven 3 und 4 repräsentieren demgegenüber mehr den ökologischen Normalfall, d.h. eine Mischung aus wenigen dominanten Arten, einigen (nicht zu) seltenen Arten und einer Reihe von "gewöhnlichen" Arten (engl.: *common species*). Solche Parameterkonstellationen führen in der Regel zu einem näherungsweise linearen Verlauf der Art-Areal-Kurve im mittleren Bereich der Flächengrößen.

Kurve 5 zeigt den Verlauf der Art-Areal-Kurve für eine homogene Lebensgemeinschaft aus gleich stark vertretenen Arten, der dem Verlauf der Art-Areal-Kurve im gewöhnlichen Koordinatensystem ähnelt. Man beachte, daß hier die "Anfangssteigung" exakt $y'(0) = 0.3/(e^{0.3} - 1) = 0.857$ beträgt, was mit der Näherung $1 - 0.3/2 = 0.850$ recht gut übereinstimmt.

Kurve 6 zeigt schließlich einen Verlauf in der extremen Situation, daß die Lebensgemeinschaft aus einer einzigen stark dominanten Art sowie mehreren seltenen Arten besteht.

Es sollte abschließend zu diesem Punkt bemerkt werden, daß die bisher durchgeführten Rechnungen alle auf der Annahme basieren, daß die Individuen der einzelnen Arten sich – im statistischen Sinn – gleichmäßig, d.h. "rein zufällig" über der Fläche verteilen (Poisson-Modell). Dies ist in konkreten Anwendungen aber durchaus nicht immer der Fall; insofern können Art-Areal-Kurven, die auf empirischen Studien beruhen, unterschiedlich ausfallen, je nach der räumlichen Lage des Areals mit der kleinsten Bezugsfläche.

Eine dritte Art von Kurven, die in der Ökologie von großer substantieller Bedeutung sind, bilden die sogenannten *Art-Abundanz-Kurven*. Hierzu betrachtet man die mittlere Anzahl $\psi(k) = E(S(k))$ der Arten, die in einem Gebiet A fester Größe $x = m(A)$ durch genau k Individuen repräsentiert werden, wobei $k = 0, 1, 2, \dots$. Für $k = 0$ erhält man gerade die mittlere Anzahl der *nicht erfaßten* ("sehr seltenen") Arten, für $k = 1$ die mittlere Anzahl der Arten, die nur mit einem einzigen Individuum vertreten sind ("seltene" Arten), usw. Art-Abundanz-Kurven können also in gewisser Weise als gemeinsame Verallgemeinerungen von Art-Repräsentanz- und Art-Areal-Kurven aufgefaßt werden. Bezeichnet wieder N_i die Anzahl der von Art i vertretenden Individuen und

$$Z_i(k) = \begin{cases} 1, & \text{falls } N_i = k \\ 0, & \text{sonst,} \end{cases} \quad i = 1, \dots, s,$$

so läßt sich $S(k)$ darstellen vermöge

$$S(k) = \sum_{i=1}^s Z_i(k)$$

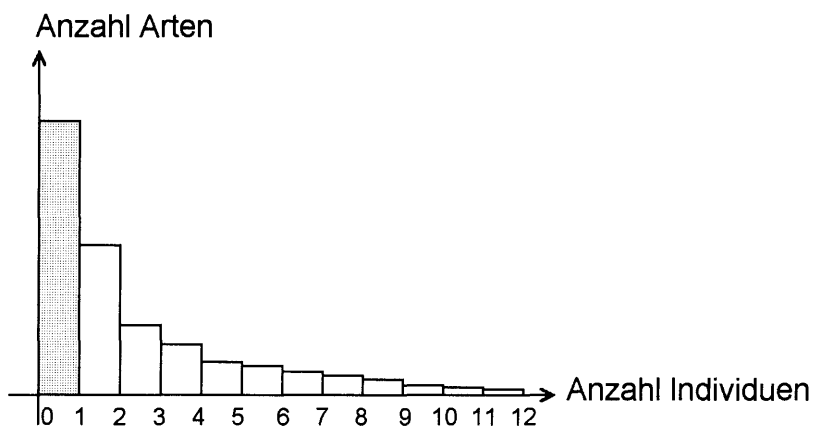
und damit

$$\psi(k) = E\left(\sum_{i=1}^s Z_i(k)\right) = \sum_{i=1}^s E(Z_i(k)) = \sum_{i=1}^s P(N_i = k), \quad k = 0, 1, 2, \dots$$

In einem Poisson-Modell entspricht dies dem Ausdruck

$$\psi(k) = \sum_{i=1}^s e^{-\lambda_i x} \frac{(\lambda_i x)^k}{k!}, \quad k = 0, 1, 2, \dots$$

für ein Gebiet A der Größe $x > 0$. Art-Abundanz-Kurven lassen sich natürlich für beliebige Verteilungsmodelle, insbesondere den "deterministischen" Fall fester Populationen, bestimmen. Dazu gehören auch empirische Art-Abundanz-Kurven für konkrete Lebensgemeinschaften (vgl. Pielou (1977), S. 271, Abb. 18.1.).



Art-Abundanz-Kurve

Graphisch lassen sich Art-Abundanz-Kurven auch in Form eines Histogramms darstellen, wobei die zum Wert $k = 0$ gehörige Klasse in der Regel nicht angegeben werden kann, da die zugehörigen Arten i.a. unbekannt sind.

Als Gesamtfläche unter der – nicht normierten – Art-Abundanz-Kurve ergibt sich

$$\sum_{k=0}^{\infty} \psi(k) = \sum_{k=0}^{\infty} \sum_{i=1}^s P(N_i = k) = \sum_{i=1}^s \sum_{k=0}^{\infty} P(N_i = k) = \sum_{i=1}^s 1 = s,$$

so daß im zugehörigen Histogramm die Säulenhöhen gerade

$$\psi^*(k) = \frac{\psi(k)}{s}, \quad k = 0, 1, 2, \dots$$

betragen.

Anhand der Art-Abundanz-Kurve kann auch die Wahrscheinlichkeit dafür berechnet werden, daß eine – zufällig ausgewählte – Art im Untersuchungsgebiet mit k Individuen vertreten ist. Ist nämlich I eine über der Menge $\{1, 2, \dots, s\}$ gleichverteilte Zufallsvariable, die die ausgewählte Art mit der Abundanz N_I angibt, so folgt

$$\begin{aligned} P(N_I = k) &= \sum_{i=1}^s P(N_I = k \mid I = i)P(I = i) = \frac{1}{s} \sum_{i=1}^s P(N_i = k) \\ &= \frac{\psi(k)}{s} = \psi^*(k), \quad k = 0, 1, 2, \dots \end{aligned}$$

Insofern entsprechen die Histogrammflächen gerade den Wahrscheinlichkeiten dafür, daß eine – zufällig ausgewählte – Art mit einer gegebenen Zahl von Individuen in dem Gebiet vertreten ist.

Ist die Zahl s der möglichen Arten groß, so kann die Art-Abundanz-Kurve auch als *Mischung* von Verteilungen interpretiert werden (vgl. Pielou (1977), S. 271f.). Für das Poisson-Modell kann man sich etwa vorstellen, daß die Werte λ_i Realisierungen einer reellen, positiven Zufallsgrößen Λ mit einer Verteilungsdichte f sind, so daß nach dem Gesetz der großen Zahlen bei unbegrenzter Artenzahl folgt

$$\begin{aligned} \Psi(k) &:= \lim_{s \rightarrow \infty} \frac{\psi(k)}{s} = \lim_{s \rightarrow \infty} \frac{1}{s} \sum_{i=1}^s e^{-\lambda_i x} \frac{(\lambda_i x)^k}{k!} = E \left(e^{-\Lambda x} \frac{(\Lambda x)^k}{k!} \right) \\ &= \int_0^{\infty} e^{-\lambda x} \frac{(\lambda x)^k}{k!} f(\lambda) d\lambda, \quad k = 0, 1, 2, \dots \end{aligned}$$

Wir wollen im folgenden einmal genauer untersuchen, welche Verteilung sich für Ψ ergibt, wenn f die Dichte einer *Gamma-Verteilung* ist, d.h. wenn gilt

$$f(\lambda) = f_{\gamma, \nu}(\lambda) = \frac{\nu^\gamma}{\Gamma(\gamma)} \lambda^{\gamma-1} e^{-\nu\lambda}, \quad \lambda > 0$$

mit Parametern $\gamma, \nu > 0$. Dabei bezeichnet Γ wieder die Euler'sche Gamma-Funktion (siehe S. 32). Für ganzzahlige Werte von γ stimmt f übrigens überein mit der Dichte der Erlang-Verteilung, die im Zusammenhang mit den Abstandsverfahren schon früher behandelt wurde (S. 34f.). Die Gamma-Verteilung besitzt den Erwartungswert $\mu = \gamma/\nu$ und die Varianz $\sigma^2 = \gamma/\nu^2$, wodurch die Bedeutung der Parameter klar wird, wenn man diese Gleichungen nach γ und ν auflöst:

$$\gamma = \frac{\mu^2}{\sigma^2}, \quad \nu = \frac{\mu}{\sigma^2}.$$

Für große Werte von γ unterscheidet sich nach dem zentralen Grenzwertsatz die Dichte $f_{\gamma, \nu}$ von der Dichte der Normalverteilung mit Erwartungswert μ und Varianz σ^2 immer weniger.

Für die durch Ψ gegebene Verteilung erhält man jetzt durch Integration

$$\begin{aligned} \Psi(k) &= \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^k}{k!} \frac{\nu^\gamma}{\Gamma(\gamma)} \lambda^{\gamma-1} e^{-\nu\lambda} d\lambda = \frac{\nu^\gamma x^k}{\Gamma(\gamma)k!} \int_0^\infty e^{-(x+\nu)\lambda} \lambda^{\gamma+k-1} d\lambda \\ &= \frac{\nu^\gamma x^k}{\Gamma(\gamma)k!} \frac{\Gamma(\gamma+k)}{(x+\nu)^\gamma} = \begin{cases} p^\gamma, & k=0 \\ \frac{\Gamma(\gamma+k)}{\Gamma(\gamma)\Gamma(k)} \frac{p^\gamma q^k}{k} = B(\gamma, k) \frac{p^\gamma q^k}{k}, & k=1, 2, \dots \end{cases} \end{aligned}$$

mit

$$p := \frac{\nu}{x+\nu}, \quad q := 1-p = \frac{x}{x+\nu} \quad \text{und} \quad B(\gamma, k) := \frac{\Gamma(\gamma+k)}{\Gamma(\gamma)\Gamma(k)}$$

(sog. *Beta-Funktion*). Ist $\gamma = m \in \mathbb{N}$ ganzzahlig, also f die Dichte einer Erlang-Verteilung, so ergibt sich für Ψ gerade eine *negative Binomialverteilung*:

$$\Psi(k) = \binom{m+k-1}{k} p^m q^k, \quad k=0, 1, 2, \dots$$

und für den speziellen Fall $m=1$ eine *geometrische Verteilung*:

$$\Psi(k) = pq^k, \quad k=0, 1, 2, \dots$$

Es gibt zahlreiche numerische Beispiele ökologischer Untersuchungen, bei denen die Art-Abundanz-Kurven einen der geometrischen oder der negativen Binomialverteilung ähnelnden Verlauf haben (vgl. Pielou (1977), S. 278ff.), was bis zu einem gewissen Grad das Mischungsmodell auch praktisch rechtfertigt. Der Parameter ν kann dabei als eine für die jeweilige Lebensgemeinschaft charakteristische Größe angesehen werden, der als "Formparameter" den Grad des Abnehmens der Wahrscheinlichkeiten $\Psi(k)$ für große k steuert.

Wir kommen zum Abschluß dieses Kapitels noch einmal auf das Problem zurück, daß das zur Art-Abundanz-Kurve gehörige Histogramm nicht exakt bestimmt werden kann, wenn die mittlere Anzahl $\psi(0)$ der in dem Untersuchungsgebiet nicht

erfaßten Arten unbekannt ist. In diesem Fall macht es aber Sinn, das *bedingte* Histogramm (für Abundanzen $k \geq 1$) zu betrachten, d.h. die bedingten Wahrscheinlichkeiten

$$P(N_I = k \mid N_I \geq 1) = \frac{P(N_I = k)}{P(N_I \geq 1)} = \frac{\psi^*(k)}{1 - \psi^*(0)}, \quad k = 1, 2, 3, \dots$$

bzw. für den Grenzfall einer unendlich großen Artenzahl im Mischungsmodell

$$\frac{\Psi(k)}{1 - \Psi(0)} = B(\gamma, k) \frac{p^\gamma q^k}{(1 - p^\gamma)^k}, \quad k = 1, 2, 3, \dots$$

Betrachtet man hier wiederum den Grenzfall $\gamma \rightarrow 0$, was der Situation entspricht, daß die Lebensgemeinschaft aus einer sehr großen Zahl seltener Arten besteht, so erhält man im letzten Fall

$$\begin{aligned} \Psi^*(k) &:= \lim_{\gamma \rightarrow 0} \frac{\Psi(k)}{1 - \Psi(0)} = \lim_{\gamma \rightarrow 0} \frac{\Gamma(\gamma + k)}{\Gamma(k)} \frac{p^\gamma q^k}{k} \frac{1}{\Gamma(\gamma)(1 - p^\gamma)} \\ &= \frac{q^k}{k} \lim_{\gamma \rightarrow 0} \frac{1}{\Gamma(\gamma)(1 - p^\gamma)}, \quad k = 1, 2, 3, \dots \end{aligned}$$

mit

$$\lim_{\gamma \rightarrow 0} \Gamma(\gamma)(1 - p^\gamma) = \lim_{\gamma \rightarrow 0} \gamma \Gamma(\gamma) \frac{1 - p^\gamma}{\gamma} = \left\{ \lim_{\gamma \rightarrow 0} \Gamma(\gamma + 1) \right\} \left\{ -\frac{d}{d\gamma} p^\gamma \Big|_{\gamma=0} \right\} = -\ln p.$$

Hierbei wurde von der Stetigkeit der Gamma-Funktion sowie der Funktionalgleichung¹

$$z \cdot \Gamma(z) = \Gamma(z + 1), \quad z > 0$$

Gebrauch gemacht. Die asymptotische bedingte Verteilung der Anzahl der durch k Individuen repräsentierten Arten für $k \geq 1$ erhält damit abschließend die Form

$$\Psi^*(k) = \beta \frac{q^k}{k}, \quad k = 1, 2, 3, \dots$$

mit dem Normierungsfaktor

$$\beta = -\frac{1}{\ln(1 - q)} = -\frac{1}{\ln p} = \frac{1}{\ln(1/p)}.$$

¹ Mit partieller Integration erhält man

$$\Gamma(z + 1) = \int_0^\infty t^z e^{-t} dt = -t^z e^{-t} \Big|_0^\infty + \int_0^\infty z t^{z-1} e^{-t} dt = z \Gamma(z)$$

Diese Verteilung heißt auch *logarithmische Verteilung* (engl.: *log-series distribution*), weil die Terme q^k/k für $k \geq 1$ die Summanden der Taylor-Entwicklung der Funktion $q \mapsto -\ln(1 - q)$ darstellen.

Wegen $p = \nu/(x + \nu)$ hängt der Faktor β noch von der Flächengröße $x = m(A)$ ab, und zwar gemäß

$$\beta = \frac{1}{\ln(1 + x/\nu)}.$$

Führt man die obigen Grenzübergänge $s \rightarrow \infty$ und $\gamma \rightarrow 0$ nicht nacheinander, sondern gleichzeitig durch, d.h. wählt man $\gamma = \gamma(s)$ geeignet in Abhängigkeit von s , so kann man die mittlere Anzahl $E(S^*)$ der vorhandenen Arten nach oben beschränken, was der Realität sicher näherkommt. Dazu betrachten wir die Zufallsgrößen

$$\bar{Z}_i(0) = 1 - Z_i(0), \quad i = 1, 2, \dots, s,$$

die angeben, ob die Art i vorhanden ist ($\bar{Z}_i(0) = 1$) oder nicht ($\bar{Z}_i(0) = 0$). Es ist dann die Anzahl der vorhandenen Arten

$$S^* = \sum_{i=1}^s \bar{Z}_i(0)$$

eine Summe unabhängiger Zufallsvariablen mit "kleiner" Erfolgswahrscheinlichkeit

$$E(\bar{Z}_i(0)) = 1 - e^{-\lambda_i x}, \quad \text{also} \quad E(S^*) = \sum_{i=1}^s (1 - e^{-\lambda_i x}) \approx s(1 - p^{\gamma(s)}).$$

Strebt diese Größe für $s \rightarrow \infty$ einem endlichen Wert s^* zu, besitzt S^* nach dem Gesetz der seltenen Ereignisse asymptotisch eine Poisson-Verteilung¹:

$$P(S^* = k) \approx e^{-s^*} \frac{s^{*k}}{k!}, \quad k = 0, 1, \dots$$

Hierauf kommen wir weiter unten noch einmal zurück.

Die mittlere Anzahl M der Individuen, mit der sich eine – unter den vertretenen "zufällig" gewählte – Art im Gebiet A manifestiert, ergibt sich unter der logarithmischen Verteilung zu

$$M = \sum_{k=1}^{\infty} k \beta \frac{q^k}{k} = \beta \frac{q}{p} = \beta \left(\frac{1}{p} - 1 \right) = \frac{x}{\nu \ln(1 + x/\nu)},$$

was bei festgehaltenem Parameter ν erwartungsgemäß monoton mit der Flächengröße x wächst, und zwar schwach sublinear.

¹ Wir benötigen hier eine etwas allgemeinere Form des einfachen Poisson'schen Gesetzes, wie es z.B. in Krengel (1991), Kapitel 5.4 beschrieben ist. Vgl. auch Hartung et al. (1989), Kapitel II,9.

Aus der letzten Gleichung wird auch die Bedeutung des Parameters ν deutlich: aufgrund bekannter Abschätzungen für die Logarithmusfunktion¹ folgt etwa

$$1 \leq M \leq 1 + \frac{x}{2\nu} \quad \text{mit} \quad M \approx 1 + \frac{x}{2\nu} \quad \text{für} \quad x \ll \nu,$$

d.h. für "kleine" Flächengrößen x ist ν näherungsweise umgekehrt proportional zu $M - 1$, wenn M nahe bei 1 liegt.

Der Parameter β wird in der ökologischen Literatur üblicherweise als Maß der "Diversität" einer Lebensgemeinschaft interpretiert, wenn die bedingte Art-Abundanz-Kurve Ψ^* – zumindest näherungsweise – durch eine logarithmische Verteilung beschreibbar ist. Dies läßt sich leicht nachvollziehen, wenn man an die Definition von β denkt: wegen

$$\Psi^*(k) = \beta \frac{q^k}{k}, \quad k = 1, 2, 3, \dots$$

ist ja

$$\Psi^*(1) = \beta q = \beta \left(1 - e^{-1/\beta}\right) \approx 1 - \frac{1}{2\beta + 1}$$

die Wahrscheinlichkeit dafür, daß eine im Gebiet A vorkommende Art mit genau einem Individuum vertreten ist. Je größer β und damit auch diese Wahrscheinlichkeit ausfällt, umso "gleichmäßiger" (und zugleich seltener) sind die vertretenen Arten also im Gebiet A repräsentiert. Ist β und damit auch $\Psi^*(1)$ dagegen klein, werden häufiger Arten mit höheren Individuenzahlen repräsentiert, d.h. das Entstehen dominanter Arten wird begünstigt. Dies erklärt u.a. auch, warum bei logarithmischer Verteilung die "Diversität" abnimmt, wenn die Flächengröße x zunimmt, da in diesem Fall ja auch die mittlere Zahl M der Individuen je vorhandener Art und damit die Möglichkeit der "Ungleichverteilung" der Abundanzen wächst.

Es sollte noch darauf hingewiesen werden, daß in der ökologischen Literatur die "Diversität" üblicherweise nicht durch die Größe β selbst, sondern durch die Größe

$$\alpha := s^* \beta$$

charakterisiert wird, wobei s^* wieder die mittlere im Gebiet A vorhandene Zahl von Arten bezeichnet (vgl. Pielou (1977), S. 273ff., Greig-Smith (1983), S. 159ff. und Krebs (1985), S. 514ff.). Bei logarithmischer Verteilung ergibt sich damit noch der folgende Zusammenhang zwischen der mittleren Anzahl N der insgesamt im Gebiet A vertretenen Individuen und der mittleren Anzahl der vorhandenen Arten s^* :

$$s^* = \alpha \ln \left(1 + \frac{N}{\alpha}\right).$$

Wegen

$$N = s^* M = s^* \beta \left(\frac{1}{p} - 1\right) = \alpha \left(\frac{1}{p} - 1\right)$$

¹ Es gilt z.B. $\frac{z}{1+z/2} \leq \ln(1+z) \leq z, \quad z \geq 0.$

ist nämlich

$$\frac{1}{p} = 1 + \frac{N}{\alpha} \quad \text{bzw.} \quad \frac{1}{\beta} = \ln(1/p) = \ln\left(1 + \frac{N}{\alpha}\right)$$

und damit

$$s^* = \frac{\alpha}{\beta} = \alpha \ln\left(1 + \frac{N}{\alpha}\right),$$

wie behauptet.

Vergrößert man das Gebiet A , so wächst die mittlere Zahl s^* der vorhandenen Arten mit dem Logarithmus der Flächengröße x , wenn die "Diversität" α konstant bleibt, da sich aus der obigen Rechnung als Nebenergebnis auch noch

$$\frac{N}{\alpha} = \frac{x}{\nu}, \quad \text{also} \quad N = \frac{\alpha x}{\nu},$$

d.h.

$$s^* = \alpha \ln\left(1 + \frac{x}{\nu}\right)$$

ergibt. Zeichnet man die entsprechende Kurve in einer doppelt-logarithmischen Skala auf, d.h. betrachtet man die Funktion (vgl. S. 56)

$$y = \log_a \alpha + \log_a \left(\ln \left(1 + \frac{a^x}{\nu} \right) \right),$$

so erscheint die entsprechende Art-Areal-Kurve bereichsweise annähernd linear, was die früheren Überlegungen auch noch einmal aus anderer Sicht rechtfertigt. Im Gegensatz zu der Größe β , die in jedem Fall abhängig von der Flächengröße x ist, kann die Größe α also in gewissen Modellen als näherungsweise unabhängig von der Flächengröße angesehen werden, was aus Sicht der Ökologie sicher wünschenswert ist.

Wir kommen zum Abschluß dieses Abschnitts noch einmal kurz auf den oben angesprochenen gleichzeitigen Grenzübergang $s \rightarrow \infty$ und $\gamma(s) \rightarrow 0$ zurück. Man kann dann die zuletzt angestellten Mittelwertbetrachtungen nämlich sofort auch auf Verteilungseigenschaften ausdehnen. Bezeichnet wieder S^* die Anzahl der im Gebiet A vertretenen Arten und sind Y_i , $i = 1, 2, \dots$ unabhängige Zufallsvariablen mit jeweils logarithmischer Verteilung (entsprechend der realisierten Anzahl von Individuen), die auch von S^* unabhängig sind, so wird durch

$$X = \sum_{i=1}^{S^*} Y_i$$

gerade die Gesamtzahl aller beobachteten Individuen beschrieben. Wie oben erhält man wieder

$$N = E(X) = E(S^*)E(Y) = s^*M,$$

wobei Y eine wie die Y_i logarithmisch verteilte Zufallsvariable bezeichnet; darüberhinaus läßt sich zeigen, daß X eine negative Binomialverteilung besitzt:

$$P(X = k) = \binom{s^*\beta + k - 1}{k} p^{s^*\beta} q^k, \quad k = 0, 1, \dots$$

(dies gilt auch für nicht-ganzzahlige Werte von $s^*\beta$ ¹). Der Beweis dieser Beziehung ist nicht elementar, kann aber elegant unter Verwendung *wahrscheinlichkeitserzeugender Funktionen* (vgl. Krengel (1991), §7) geführt werden. Sei dazu

$$g_Y(t) = E(t^Y) = \sum_{k=1}^{\infty} t^k P(Y = k),$$

$$g_{S^*}(t) = E(t^{S^*}) = \sum_{k=0}^{\infty} t^k P(S^* = k),$$

für $0 \leq t \leq 1$. g_Y heißt die wahrscheinlichkeitserzeugende Funktion der Verteilung von Y . Die Einzelwahrscheinlichkeiten $P(Y = k)$ werden also durch die Koeffizienten der Taylorreihe von g_Y dargestellt. Für die wahrscheinlichkeitserzeugende Funktion der Verteilung von X erhält man mit der Unabhängigkeitsannahme

$$g_X(t) = E(t^X) = \sum_{n=0}^{\infty} E(t^{\sum_{k=1}^n Y_k}) P(S^* = n)$$

$$= \sum_{n=0}^{\infty} [E(t^Y)]^n P(S^* = n) = g_{S^*}(g_Y(t)).$$

Im einzelnen ist

$$g_Y(t) = \sum_{k=1}^{\infty} t^k P(Y = k) = \beta \sum_{k=1}^{\infty} \frac{(qt)^k}{k} = -\beta \ln(1 - qt) = \frac{\ln(1 - qt)}{\ln(1 - q)},$$

$$g_{S^*}(t) = \sum_{k=0}^{\infty} t^k P(S^* = k) = e^{-s^*} \sum_{k=0}^{\infty} \frac{(s^*t)^k}{k!} = e^{-s^*} e^{s^*t} = e^{s^*(t-1)},$$

für $0 \leq t \leq 1$. Damit folgt

$$g_X(t) = g_{S^*}(g_Y(t)) = \frac{e^{-s^*}}{(1 - qt)^{s^*\beta}} = \frac{p^{s^*\beta}}{(1 - qt)^{s^*\beta}}$$

$$= \sum_{k=0}^{\infty} \binom{s^*\beta + k - 1}{k} p^{s^*\beta} q^k t^k = \sum_{k=0}^{\infty} P(X = k) t^k,$$

für $0 \leq t \leq 1$, woraus durch Koeffizientenvergleich die angegebene negative Binomialverteilung für X folgt.

¹ Für $a \in \mathbb{R}$ ist $\binom{a}{k} = \frac{a(a-1)\cdots(a-k+1)}{k!}$, $k = 0, 1, \dots$ mit $\binom{a}{0} = 1$.

Diversitätsindices

Ein wichtiger Aspekt der numerischen Struktur von Lebensgemeinschaften wird vollständig vernachlässigt, wenn die Zusammensetzung der Lebensgemeinschaft nur durch die Anzahl vorhandener Arten beschrieben wird. Dieses Vorgehen vernachlässigt die Information, daß einige Arten selten und andere häufig sind. Intuitiv erscheint eine Lebensgemeinschaft mit sieben gleich häufigen Arten diverser als eine mit sieben Arten, bei der aber 40% zu der häufigsten Art gehören und nur 5% zu jeder der drei seltensten. ... Trotzdem hat jede Lebensgemeinschaft den gleichen Artenreichtum.

Wenn die Lebensgemeinschaft klar definiert ist (z.B. eine Lebensgemeinschaft von Singvögeln in einem Waldgebiet), können Zählungen der Anzahl der Individuen jeder Art für viele Zwecke genügen. Wenn wir jedoch an allen Tieren des Waldes interessiert sind, ergibt es wenig Sinn, die gleiche Art der Quantifizierung für Protozoen¹, Asseln, Vögel und Hirsche zu benutzen. Der enorme Größenunterschied bedeutet, daß Zählungen sehr irreführend wären. Wir begegnen auch großen Schwierigkeiten, wenn wir versuchen, Pflanzen oder andere modulare Organismen zu zählen. Sollen wir die Anzahl der Schosse, Blätter, Stämme oder Genets zählen? Ein Weg dieses Problem zu umgehen, ist die Beschreibung der Lebensgemeinschaft in der Form von Biomasse (oder Produktionsrate der Biomasse) pro Art und Fläche.

[aus Begon/Harper/Townsend (1991)]

Ecological communities do not all contain the same number of species, and one of the currently active areas of research in community ecology is the study of species richness or diversity. A. W. Wallace (1878) recognized that animal life was on the whole more abundant and varied in the tropics than in other parts of the globe, and the same applies to plants. Other patterns of variation have long been known on islands; small or remote islands have fewer species than large islands or those nearer continents... The regularity of these patterns for many taxonomic groups suggests that they have been produced in conformity with a set of basic principles rather than accidents of history. How can we explain these trends in species diversity?

...

The simplest measure of species diversity is to count the *number of species*. In such a count we should include only resident species, not accidental or temporary immigrants. It may not always be easy to decide which species are accidentals: Is a bottomland tree species growing on a ridge top an accidental species or a resident one? The number of species is the first and oldest concept of species diversity and is called *species richness*.

A second concept of species diversity is that of *heterogeneity*. One problem with counting the number of species as a measure of diversity is that it treats rare species and common species as equals. A community with two species might be divided in two extreme ways:

¹ Einzeller, z.B. Amöben, Geißeltierchen, Wimperntierchen

	Community 1	Community 2
Species A	99	50
Species B	1	50

The second community would seem intuitively to be more diverse than the first. Peet (1974) suggested that we combine the concepts of number of species and relative abundance into a single concept of *heterogeneity*. Heterogeneity is higher in a community when there are more species and when the species are equally abundant.

...

Several measures of heterogeneity are in use ..., and the most popular has been borrowed from **information theory**. This approach is preferred by some because it is independent of any hypothetical distribution such as the log-normal. The main objective of information theory is to try to measure the amount of *order* (or disorder) contained in a system (Margalef (1958)). Four types of information might be collected regarding order in the community: (1) the number of species, (2) the number of individuals in each species, (3) the places occupied by individuals of each species, and (4) the places occupied by individuals as separate individuals. In most community work, only data of types 1 and 2 are obtained.

Information theory, Margalef suggested, provides one way to escape some of the difficulties of the log-normal curve and the logarithmic series. We ask the question, How difficult would it be to predict correctly the species of the next individual collected? This is the same problem faced by communication engineers interested in predicting correctly the name of the next letter in a message. This uncertainty can be measured by the Shannon-Wiener function:

$$H = - \sum_{i=1}^s p_i \log_2 p_i$$

where

H = information content of sample (bits/individual) = index of species diversity

s = number of species

p_i = proportion of total sample belonging to the i -th species

Information content is a measure of the amount of uncertainty, so the larger the value of H , the greater the uncertainty. A message such as *bbbbbb* has no uncertainty in it, and $H = 0$. For our example of two species of 99 and 1 individuals,

$$\begin{aligned} H &= -[(p_1)(\log_2 p_1) + (p_2)(\log_2 p_2)] \\ &= -[(0.99)(\log_2 0.99) + (0.01)(\log_2 0.01)] = 0.081 \text{ bit/individual} \end{aligned}$$

For a sample of two species with 50 individuals in each,

$$\begin{aligned} H &= -[(0.50)(\log_2 0.50) + (0.50)(\log_2 0.50)] \\ &= 1.00 \text{ bit/individual} \end{aligned}$$

This agrees with our intuitive feeling that the second sample is more diverse than the first sample.

[aus Krebs (1985)]

There are three reasons why ecologists are interested in ecological diversity and its measurement. First, despite changing fashions and preoccupations, diversity has remained a central theme in ecology. The well documented patterns of spatial and temporal variation in diversity which intrigued the early investigators of the natural world . . . continue to stimulate the minds of ecologists today Second, measures of diversity are frequently seen as indicators of the wellbeing of ecological systems. Thirdly, considerable debate surrounds the measurement of diversity. Diversity may appear to be a straightforward concept which can be quickly and painlessly measured. This is because most people have a ready intuitive grasp of what is meant by diversity and have little difficulty in accepting, say, that tropical rain forests are more diverse than temperate woodlands or that there is a high diversity of organisms in coral reefs. Yet diversity is rather like an optical illusion. The more it is looked at, the less clearly defined it appears to be and viewing it from different angles can lead to different perceptions of what is involved. The problem has been exacerbated by the fact that ecologists have devised a huge range of indices and models for measuring diversity. Despite, or perhaps as a result of these, diversity has a knack of eluding definition and in one instance Hurlbert . . . even went so far as to decry it as a 'non-concept'.

There is however a simple explanation why diversity is so hard to define. That is because diversity consists of not one but two components. These are first the variety and secondly the relative abundance of species. . . . It is sufficient for now to note that diversity can be measured by recording the number of species, by describing their relative abundances or by using a measure which combines the two components. It is important that ecologists should understand how to measure diversity and what they mean by it. Diversity lies at the root of some of the most fundamental and exciting questions in theoretical and applied ecology.

...

It is rarely feasible, or desirable, to census every individual in a community. Such a strategy would be prohibitively time-consuming and expensive; it would also damage or possibly even destroy the community in question. Ecologists therefore rely on **sampling** to provide an accurate picture of community composition. A great deal of effort over past decades has been devoted to making **sampling techniques** as efficient as possible. . . . Diversity studies raise a number of special problems where sampling is concerned. For example can individuals be sampled **randomly**? What size should samples be? What happens if individuals are not easily recognizable?

[aus Magurran (1988)]

Wir werden in diesem Abschnitt nur die sogenannte α -Diversität einer Lebensgemeinschaft betrachten, die sich auf den Artenreichtum innerhalb eines Habitats bezieht (im Gegensatz zur β -Diversität, die den Grad des Wechsels in der Artzusammensetzung z.B. entlang eines Umweltgradienten beschreibt). Unter der "Diversität" einer Lebensgemeinschaft, die aus s potentiellen Arten besteht, soll dabei in Anlehnung an die einleitenden Ausführungen mathematisch eine Funktion D verstanden werden, die von den Anteilen $\mathbf{p} = (p_1, \dots, p_s)$ abhängt, mit denen die s Arten in der Lebensgemeinschaft vertreten sind. Der Definitionsbereich \mathcal{D} von D ist damit gegeben durch

$$\mathcal{D} = \left\{ \mathbf{p} \mid \mathbf{p} = (p_1, \dots, p_s), p_i \geq 0, i = 1, \dots, s, \sum_{i=1}^s p_i = 1; s \in \mathbb{N} \right\}.$$

Es gibt einige mathematische Forderungen an die Funktion D , die wesentliche Eigenschaften dessen widerspiegeln, was man aus Sicht der Ökologie unter "Diversität" verstehen will:

- (D1) $D(\mathbf{p}) \geq 0$. Diese Eigenschaft dient zur Festlegung einer sinnvollen Skala für die Werte von D .
- (D2) $D(\mathbf{p}) = 0$ genau dann, wenn die Lebensgemeinschaft nur aus einer Art besteht (d.h. $p_i = 1$ für ein i).
- (D3) $D(\mathbf{p}) = D(p_{i_1}, \dots, p_{i_s})$ für alle Permutationen (i_1, \dots, i_s) der Zahlen $1, \dots, s$. Diese Eigenschaft besagt, daß die Diversität nicht von der Art der Benennung oder Numerierung der Arten abhängt.
- (D4) $D(\mathbf{p}) \leq D(1/s, \dots, 1/s)$. Diese Eigenschaft besagt, daß die Diversität der Lebensgemeinschaft am höchsten ist, wenn alle potentiellen Arten gleich häufig vertreten sind.
- (D5) $D(\mathbf{p}, 0) = D(\mathbf{p})$. Diese Eigenschaft bedeutet, daß die Diversität gleichbleibt, wenn die Lebensgemeinschaft um eine irrelevante Art (mit Anteil 0) erweitert wird. $(\mathbf{p}, 0)$ entspricht dabei den Anteilen der ursprünglichen Arten in der Lebensgemeinschaft.
- (D6) $D(1/(s-1), \dots, 1/(s-1), 0) \leq D(1/s, \dots, 1/s)$ für $s > 1$. Diese Eigenschaft bedeutet, daß bei Entfernung einer Art aus einer homogenen Lebensgemeinschaft (d.h. alle Artanteile sind gleich groß: $p_i = 1/s$) die Diversität abnimmt.

Leider legen die Forderungen (D1) bis (D6) "die" Funktion D noch nicht eindeutig fest. Dies kann man schon daran erkennen, daß mit D auch $T(D)$ ein Diversitätsindex mit denselben Eigenschaften ist für jede strikt monoton wachsende Abbildung $T : [0, \infty) \rightarrow [0, \infty)$ mit $T(0) = 0$.

Beispiele für (wesentlich) unterschiedliche Diversitätsindices D , die alle sechs Forderungen erfüllen, sind etwa:

$$D_E(\mathbf{p}) = - \sum_{i=1}^s p_i \cdot \log_a p_i \quad (\text{Shannon-Wiener-Entropie})$$

mit der Festlegung $0 \cdot \log_a 0 = 0$, wobei für die Basis a des Logarithmus (wegen (D1)) gelte: $a > 1$;

$$D_S(\mathbf{p}) = 1 - \left(\sum_{i=1}^s p_i^\alpha \right)^\beta \quad (\text{Simpson-Index})$$

mit $\alpha > 1$, $\beta > 0$. Der Simpson-Index wurde ursprünglich mit den Parametern $\alpha = 2$, $\beta = 1$ eingeführt und erst später durch McIntosh und Hill verallgemeinert (siehe Greig-Smith (1983), S. 163ff.). Es gibt ihn auch in Quotientenform (siehe Mühlenberg (1993), S. 358 oder Begon, Harper und Townsend (1991), S. 683):

$$D_H(\mathbf{p}) = \left(\sum_{i=1}^s p_i^\alpha \right)^{-\beta} \quad (\text{Hill-Index})$$

mit $\alpha > 1$, $\beta > 0$. Der Hill-Index verletzt allerdings die Forderung (D2), da er nicht den Wert 0, sondern minimal nur den Wert 1 annehmen kann. Der logarithmisch transformierte Index

$$D_L(\mathbf{p}) := \log_a (D_H(\mathbf{p}))$$

zur Basis $a > 1$ erfüllt dagegen auch Eigenschaft (D2), entsprechend der modifizierte Index

$$D_M(\mathbf{p}) = D_H(\mathbf{p}) - 1.$$

Der Nachweis der Eigenschaften (D1) bis (D3) und (D5) für die genannten Indices ist nicht schwer. Zum Nachweis von (D4) betrachte man die entsprechende Lagrange-Funktion

$$L(\mathbf{p}, \lambda) = D(\mathbf{p}) + \lambda(\mathbf{p}^{tr} \mathbf{1} - 1)$$

mit der notwendigen Bedingung

$$\nabla_{\mathbf{p}} L(\mathbf{p}, \lambda) = \nabla_{\mathbf{p}} D(\mathbf{p}) + \lambda \mathbf{1} = 0, \quad \text{also} \quad \nabla_{\mathbf{p}} D(\mathbf{p}) = \text{const}$$

für ein relatives Maximum bei \mathbf{p} . Beispielsweise gilt (Logarithmen und Potenzen jeweils komponentenweise, als Spaltenvektoren):

$$\begin{aligned} \nabla_{\mathbf{p}} D_E(\mathbf{p}) &= -\log_a(\mathbf{p}^{tr}) - \frac{1}{\ln a} \mathbf{1} \\ \nabla_{\mathbf{p}} D_S(\mathbf{p}) &= -\alpha\beta \left(D_H(\mathbf{p}) \right)^{-1+1/\beta} (\mathbf{p}^{tr})^{\alpha-1} \\ \nabla_{\mathbf{p}} D_L(\mathbf{p}) &= -\frac{\alpha\beta}{\ln a} D_H(\mathbf{p})^{1/\beta} (\mathbf{p}^{tr})^{\alpha-1} \\ \nabla_{\mathbf{p}} D_M(\mathbf{p}) &= -\alpha\beta D_H(\mathbf{p})^{1+1/\beta} (\mathbf{p}^{tr})^{\alpha-1} \end{aligned}$$

Die genannte notwendige Bedingung liefert hier in allen Fällen $\mathbf{p} = \text{const}$ und damit wegen der Nebenbedingung $\mathbf{p} = (1/s, \dots, 1/s)$ als Maximumstelle.

Zum Nachweis von (D6) beachte man

$$D_E(1/(s-1), \dots, 1/(s-1), 0) = \log_a(s-1) < \log_a s = D_E(1/s, \dots, 1/s)$$

sowie

$$\sum_{i=1}^{s-1} \left(\frac{1}{s-1}\right)^\alpha = \frac{1}{(s-1)^{\alpha-1}} > \frac{1}{s^{\alpha-1}} = \sum_{i=1}^s \left(\frac{1}{s}\right)^\alpha,$$

woraus die Behauptung für den Simpson- und modifizierten Hill-Index folgt.

Aus den bisher genannten Eigenschaften folgt in der Regel nicht, daß bei Entfernung einer Art aus einer (nicht-homogenen) Lebensgemeinschaft die Diversität grundsätzlich abnimmt. Das folgende Beispiel zeigt, daß sowohl eine Abnahme als auch eine Zunahme der Diversität bei Entfernung einer Art möglich ist (Entropie D_E zur Basis $a = 10$, Simpson-Index D_S mit $\alpha = 2$, $\beta = 1$):

D_E	D_S	Artanteile					
0.650	0.700	0.1	0.1	0.1	0.1	0.1	0.5
0.698	0.800	0.2	0.2	0.2	0.2	0.2	0.0
0.565	0.641	0.0	0.1	0.1	0.1	0.1	0.5

In diesem Beispiel führt die Entfernung einer stark dominanten Art bei gleichzeitigem Vorliegen vieler "seltener" Arten zu einer Erhöhung der Diversität, während die Entfernung einer "seltenen" Art zur Abnahme der Diversität führt.

In der Praxis lassen sich Diversitäten meist nur auf der Basis einer *Stichprobe* ermitteln, indem aus der Lebensgemeinschaft eine gewisse Zahl n von Individuen entnommen wird. Bezeichnen wir mit X_i die zugehörigen beobachteten Abundanzen für die Art i , so bilden die Quotienten $\hat{p}_i = \frac{X_i}{n}$ für $i = 1, 2, \dots, s$ Schätzgrößen für die "wahren" Artanteile $p_i = \frac{N_i}{N}$, wenn die Lebensgemeinschaft aus insgesamt N Individuen besteht und jede Art i mit N_i Individuen repräsentiert ist. Als empirische Diversität erhält man somit den Ausdruck $D(\hat{\mathbf{p}})$ mit $\hat{\mathbf{p}} = (\hat{p}_1, \dots, \hat{p}_s)$. $D(\hat{\mathbf{p}})$ ist damit eine Zufallsgröße, die um die "wahre" Diversität $D(\mathbf{p})$ schwankt. Wir

wollen im folgenden einige statistische Eigenschaften solcher empirisch gewonnenen Diversitäten untersuchen.

Als gemeinsame Verteilung der Abundanzen X_1, \dots, X_s erhalten wir zunächst eine sogenannte *multi-hypergeometrische Verteilung*:

$$P(X_1 = n_1, \dots, X_s = n_s) = \frac{\binom{N_1}{n_1} \binom{N_2}{n_2} \dots \binom{N_s}{n_s}}{\binom{N}{n}}$$

für $n_i = 0, \dots, m_i := \min\{N_i, n\}$, $n = n_1 + \dots + n_s$. Dies folgt aus der Tatsache, daß auf jeweils genau $\binom{N_i}{n_i}$ verschiedene Weisen n_i Individuen der Art i aus der Gesamtheit von N_i Individuen dieser Art ausgewählt werden können und jede Auswahl als gleichwahrscheinlich angenommen wird. Insbesondere ist jedes X_i selbst hypergeometrisch verteilt mit

$$P(X_i = n_i) = \frac{\binom{N_i}{n_i} \binom{N-N_i}{n-n_i}}{\binom{N}{n}}, \quad n_i = 0, \dots, m_i$$

(vgl. S. 6). Wegen $E(X_i) = \frac{N_i}{N}n$ folgt also

$$E(\hat{p}_i) = \frac{N_i}{N} = p_i, \quad i = 1, \dots, s,$$

d.h. $\hat{\mathbf{p}}$ ist ein *erwartungstreuer* Schätzer für \mathbf{p} . Für die zugehörigen Kovarianzen ergibt sich analog der Rechnung auf S. 10 für $i \neq j$ und $N \geq 2$, $n \geq 2$:

$$\begin{aligned} E(X_i X_j) &= \sum_{k=1}^s \sum_{n_k=0}^{m_k} n_i n_j P(X_1 = n_1, \dots, X_s = n_s) \\ &= \sum_{k=1}^s \sum_{n_k=0}^{m_k} n_i n_j \frac{\prod_{r=1}^s \binom{N_r}{n_r}}{\binom{N}{n}} \\ &= \sum_{n_i=1}^{m_i} \sum_{n_j=1}^{m_j} N_i N_j \binom{N_i-1}{n_i-1} \binom{N_j-1}{n_j-1} \sum_{\substack{k=1 \\ k \neq i, j}}^s \sum_{n_k=0}^{m_k} \frac{\prod_{r=1, r \neq i, j}^s \binom{N_r}{n_r}}{\binom{N}{n}} \\ &= N_i N_j \frac{\binom{N-2}{n-2}}{\binom{N}{n}} = N_i N_j \frac{n(n-1)}{N(N-1)}, \end{aligned}$$

also

$$\text{Kov}(X_i, X_j) = N_i N_j \frac{n}{N} \left(\frac{n-1}{N-1} - \frac{n}{N} \right) = -p_i p_j \frac{n(N-n)}{N-1}.$$

Für die Varianzen ergibt sich entsprechend (vgl. S. 7):

$$\sigma^2(X_i) = p_i(1-p_i) \frac{n(N-n)}{N-1}$$

bzw. für die geschätzten Artanteile:

$$\begin{aligned} \text{Cov}(\hat{p}_i, \hat{p}_j) &= -p_i p_j \frac{N-n}{n(N-1)}, \quad i \neq j, \\ \sigma^2(\hat{p}_i) &= p_i(1-p_i) \frac{N-n}{n(N-1)}. \end{aligned}$$

Die Varianz-Kovarianzmatrix Σ von $\hat{\mathbf{p}}$ lautet damit in geschlossener Form:

$$\Sigma = \frac{N-n}{n(N-1)} \begin{pmatrix} p_1 q_1 & -p_1 p_2 & -p_1 p_3 & \dots & -p_1 p_s \\ -p_1 p_2 & p_2 q_2 & -p_2 p_3 & \dots & -p_2 p_s \\ \vdots & \vdots & \vdots & & \vdots \\ -p_1 p_s & -p_2 p_s & -p_3 p_s & \dots & p_s q_s \end{pmatrix}$$

mit $q_i = 1 - p_i$, $i = 1, \dots, s$. Mit wachsendem n nähert sich Σ offensichtlich immer mehr der Nullmatrix an; dies bedeutet, daß die Variabilität in den Schätzungen \hat{p}_i immer kleiner und damit die Genauigkeit von $\hat{\mathbf{p}}$ immer größer wird. Bei einer Vollerhebung, d.h. $n = N$, stimmt $\hat{\mathbf{p}}$ sogar mit \mathbf{p} überein; in diesem Fall verschwindet der statistische Fehler ganz.

Die Situation stellt sich allerdings etwas anders dar, wenn die "wahre" Diversität $D(\mathbf{p})$ durch $D(\hat{\mathbf{p}})$ geschätzt wird. Die Erwartungstreue geht hier i.a. verloren: beispielsweise unterschätzt $D(\hat{\mathbf{p}})$ systematisch $D(\mathbf{p})$, wenn – wie im Falle der Entropie oder des Simpson-Index – die Diversität von der Form

$$D(\mathbf{p}) = \sum_{i=1}^s f(p_i)$$

ist mit einer konkaven Funktion f , etwa wenn $f''(x) < 0$ ausfällt für $0 < x < 1$. Mit der schon auf S. 32 erwähnten Jensen'schen Ungleichung erhält man nämlich in diesem Fall nur

$$E(D(\hat{\mathbf{p}})) < D(E(\hat{\mathbf{p}})) = D(\mathbf{p}).$$

Bei der Entropie zur Basis e ist z.B. $f(x) = -x \ln x$, also $f''(x) = -1/x < 0$ für $0 < x < 1$; beim einfachen Simpson-Index ist $f(x) = x(1-x)$ mit $f''(x) = -2 < 0$. In beiden Fällen ist die Schätzung $D(\hat{\mathbf{p}})$ für die Diversität also systematisch zu klein.

Um einen Eindruck von der Größe eines solchen systematischen Fehlers zu erhalten, betrachten wir die multivariate Taylor-Entwicklung der Diversität um den "wahren" Wert \mathbf{p} . Dies setzt natürlich voraus, daß die Abbildung $\mathbf{p} \mapsto D(\mathbf{p})$ hinreichend glatt, d.h. hier mindestens zweimal stetig partiell differenzierbar ist. Für sämtliche oben angegebenen Beispiele trifft dies zu. Setzen wir $\mathbf{h} = \hat{\mathbf{p}} - \mathbf{p}$, so erhalten wir zunächst approximativ

$$D(\hat{\mathbf{p}}) \approx D(\mathbf{p}) + \mathbf{h} \nabla D(\mathbf{p}) + \frac{1}{2} \mathbf{h} \nabla^2 D(\mathbf{p}) \mathbf{h}^{tr}.$$

Hierbei bezeichnet $\nabla D(\mathbf{p}) = \left(\frac{\partial}{\partial p_1} D(\mathbf{p}), \dots, \frac{\partial}{\partial p_s} D(\mathbf{p}) \right)^{tr}$ wie zuvor den Gradienten von D im Punkt \mathbf{p} und

$$\nabla^2 D(\mathbf{p}) = \begin{pmatrix} \frac{\partial^2}{\partial p_1^2} D(\mathbf{p}) & \frac{\partial^2}{\partial p_1 \partial p_2} D(\mathbf{p}) & \dots & \frac{\partial^2}{\partial p_1 \partial p_s} D(\mathbf{p}) \\ \frac{\partial^2}{\partial p_2 \partial p_1} D(\mathbf{p}) & \frac{\partial^2}{\partial p_2^2} D(\mathbf{p}) & \dots & \frac{\partial^2}{\partial p_2 \partial p_s} D(\mathbf{p}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2}{\partial p_s \partial p_1} D(\mathbf{p}) & \frac{\partial^2}{\partial p_s \partial p_2} D(\mathbf{p}) & \dots & \frac{\partial^2}{\partial p_s^2} D(\mathbf{p}) \end{pmatrix}$$

die Matrix der zweiten partiellen Ableitungen von D im Punkt \mathbf{p} . Für den Erwartungswert $E(D(\hat{\mathbf{p}}))$ ergibt sich somit approximativ

$$\begin{aligned} E(D(\hat{\mathbf{p}})) &\approx D(\mathbf{p}) + [E(\hat{\mathbf{p}} - \mathbf{p})] \nabla D(\mathbf{p}) + \frac{1}{2} E [(\hat{\mathbf{p}} - \mathbf{p}) \nabla^2 D(\mathbf{p}) (\hat{\mathbf{p}} - \mathbf{p})^{tr}] \\ &= D(\mathbf{p}) + \frac{1}{2} E [(\hat{\mathbf{p}} - \mathbf{p}) \nabla^2 D(\mathbf{p}) (\hat{\mathbf{p}} - \mathbf{p})^{tr}] \end{aligned}$$

bzw. explizit, unter Verwendung von Σ :

$$\begin{aligned} E(D(\hat{\mathbf{p}})) &\approx D(\mathbf{p}) + \frac{1}{2} E \left[\sum_{i=1}^s (\hat{p}_i - p_i)^2 \frac{\partial^2}{\partial p_i^2} D(\mathbf{p}) + \dots \right. \\ &\quad \left. + \sum_{i=1}^s \sum_{\substack{j=1 \\ i \neq j}}^s (\hat{p}_i - p_i)(\hat{p}_j - p_j) \frac{\partial^2}{\partial p_i \partial p_j} D(\mathbf{p}) \right] \\ &= D(\mathbf{p}) + \frac{1}{2} \left[\sum_{i=1}^s \sigma^2(\hat{p}_i) \frac{\partial^2}{\partial p_i^2} D(\mathbf{p}) + \sum_{i=1}^s \sum_{\substack{j=1 \\ i \neq j}}^s \text{Kov}(\hat{p}_i, \hat{p}_j) \frac{\partial^2}{\partial p_i \partial p_j} D(\mathbf{p}) \right] \\ &= D(\mathbf{p}) + \frac{N-n}{2n(N-1)} \left[\sum_{i=1}^s p_i q_i \frac{\partial^2}{\partial p_i^2} D(\mathbf{p}) - \sum_{i=1}^s \sum_{\substack{j=1 \\ i \neq j}}^s p_i p_j \frac{\partial^2}{\partial p_i \partial p_j} D(\mathbf{p}) \right] \\ &= D(\mathbf{p}) + \frac{N-n}{2n(N-1)} \left[\sum_{i=1}^s p_i \frac{\partial^2}{\partial p_i^2} D(\mathbf{p}) - \mathbf{p} \nabla^2 D(\mathbf{p}) \mathbf{p}^{tr} \right]. \end{aligned}$$

Der mittlere systematische statistische Fehler (Bias) bei der Schätzung der Diversität $D(\mathbf{p})$ durch $D(\hat{\mathbf{p}})$ liegt damit in der Größenordnung von maximal c/n mit einer Konstanten c , die nur von \mathbf{p} und den zweiten partiellen Ableitungen von D im Punkt \mathbf{p} abhängt. Durch geeignete Erhöhung des Stichprobenumfangs läßt sich dieser Fehler also quantitativ beliebig klein machen, unabhängig davon, wie groß die Lebensgemeinschaft tatsächlich ist.

Für die Abschätzung der Varianz von $D(\hat{\mathbf{p}})$ erhält man analog:

$$\begin{aligned} D^2(\hat{\mathbf{p}}) - D^2(\mathbf{p}) &= [D(\hat{\mathbf{p}}) - D(\mathbf{p})][D(\hat{\mathbf{p}}) + D(\mathbf{p})] \\ &\approx \left[\mathbf{h}\nabla D(\mathbf{p}) + \frac{1}{2}\mathbf{h}\nabla^2 D(\mathbf{p})\mathbf{h}^{tr} \right] [2D(\mathbf{p}) + \mathbf{h}\nabla D(\mathbf{p})] \\ &\approx 2(\mathbf{h}\nabla D(\mathbf{p}))D(\mathbf{p}) + (\mathbf{h}\nabla^2 D(\mathbf{p})\mathbf{h}^{tr})D(\mathbf{p}) + (\mathbf{h}\nabla D(\mathbf{p}))^2, \end{aligned}$$

also bei Berücksichtigung von Termen bis zur zweiten Ordnung in \mathbf{h} :

$$\begin{aligned} E[D^2(\hat{\mathbf{p}})] &\approx D^2(\mathbf{p}) + 2(E(\hat{\mathbf{p}} - \mathbf{p})\nabla D(\mathbf{p}))D(\mathbf{p}) + \dots \\ &\quad + E[(\hat{\mathbf{p}} - \mathbf{p})\nabla^2 D(\mathbf{p})(\hat{\mathbf{p}} - \mathbf{p})^{tr}]D(\mathbf{p}) + E[(\hat{\mathbf{p}} - \mathbf{p})\nabla D(\mathbf{p})]^2 \end{aligned}$$

bzw. nach Zusammenfassung aller Terme

$$\begin{aligned} \sigma^2(D(\hat{\mathbf{p}})) &= E(D^2(\hat{\mathbf{p}})) - E^2(D(\hat{\mathbf{p}})) \approx E[(\hat{\mathbf{p}} - \mathbf{p})\nabla D(\mathbf{p})]^2 \\ &= \sum_{i=1}^s \sigma^2(\hat{p}_i) \left(\frac{\partial}{\partial p_i} D(\mathbf{p}) \right)^2 + \sum_{\substack{i=1 \\ i \neq j}}^s \sum_{j=1}^s \text{Kov}(\hat{p}_i, \hat{p}_j) \frac{\partial}{\partial p_i} D(\mathbf{p}) \frac{\partial}{\partial p_j} D(\mathbf{p}) \\ &= \frac{N-n}{n(N-1)} \left[\sum_{i=1}^s p_i \left(\frac{\partial}{\partial p_i} D(\mathbf{p}) \right)^2 - \left(\sum_{i=1}^s p_i \frac{\partial}{\partial p_i} D(\mathbf{p}) \right)^2 \right] \\ &= \frac{N-n}{n(N-1)} V(\nabla D(\mathbf{p})), \end{aligned}$$

wobei man – wegen der formalen Analogie – den Ausdruck

$$V(\nabla D(\mathbf{p})) = \sum_{i=1}^s p_i \left(\frac{\partial}{\partial p_i} D(\mathbf{p}) \right)^2 - \left(\sum_{i=1}^s p_i \frac{\partial}{\partial p_i} D(\mathbf{p}) \right)^2$$

auch als "Varianz" von $\nabla D(\mathbf{p})$ bezeichnen kann. Man beachte dabei, daß im Punkt $\mathbf{p}_0 = (1/s, \dots, 1/s)$ als Extrempunkt von D , also $\nabla D(\mathbf{p}_0) = \mathbf{0}$, die "Varianz" $V(\nabla D(\mathbf{p}_0))$ verschwindet. Hier ist die Varianz $\sigma^2(D(\hat{\mathbf{p}}))$ der Schätzung für $D(\mathbf{p}_0)$ also von kleinerer Ordnung als c/n wie für alle übrigen Punkte $\mathbf{p} \neq \mathbf{p}_0$, wobei die Konstante c wieder im wesentlichen nur von dem Gradienten von D im Punkt \mathbf{p} abhängt.

Die Größe $V(\nabla D(\mathbf{p}))$ läßt sich natürlich ebenfalls aus der Stichprobe schätzen, und zwar analog zu oben durch

$$\hat{V}(\nabla D(\mathbf{p})) = \sum_{i=1}^s \hat{p}_i \left(\frac{\partial}{\partial p_i} D(\hat{\mathbf{p}}) \right)^2 - \left(\sum_{i=1}^s \hat{p}_i \frac{\partial}{\partial p_i} D(\hat{\mathbf{p}}) \right)^2.$$

Substituiert man in $\sigma^2(D(\hat{\mathbf{p}}))$ die Größe $\hat{V}(\nabla D(\mathbf{p}))$ für $V(\nabla D(\mathbf{p}))$, so erhält man – zumindest approximativ – eine Schätzung für die Varianz von $D(\hat{\mathbf{p}})$. Damit kann man also auch näherungsweise Konfidenzbereiche für die "wahre" Diversität $D(\mathbf{p})$ angeben (vgl. Kapitel 1: Versuchsplanung). Die Schätzung $\hat{V}(\nabla D(\mathbf{p}))$ für $V(\nabla D(\mathbf{p}))$ ist allerdings i.a. ebenfalls nicht erwartungstreu. Eine approximative Schätzung für $\sigma^2(D(\hat{\mathbf{p}}))$ ist gegeben durch

$$\hat{\sigma}^2(D(\hat{\mathbf{p}})) := \frac{N-n}{n(N-1)} \hat{V}(\nabla D(\mathbf{p})).$$

Die zuletzt gewählte vereinfachende Schreibweise läßt sich übrigens auch elegant auf die approximative Darstellung des Erwartungswerts von $D(\hat{\mathbf{p}})$ übertragen. Definieren wir analog den Differentialoperator

$$V(\nabla) = \sum_{i=1}^s p_i \frac{\partial^2}{\partial p_i^2} - \left(\sum_{i=1}^s p_i \frac{\partial}{\partial p_i} \right)^2$$

(wobei das letzte Quadrat die Hintereinanderausführung der Differentiation bezeichnet), so folgt entsprechend

$$E(D(\hat{\mathbf{p}})) \approx D(\mathbf{p}) + \frac{N-n}{2n(N-1)} [V(\nabla)] D(\mathbf{p}).$$

Eine geeignete Bias-korrigierte, approximative Schätzung für $D(\mathbf{p})$ ist damit gegeben durch

$$\hat{D}(\mathbf{p}) := D(\hat{\mathbf{p}}) - \frac{N-n}{2n(N-1)} [V(\nabla)] D(\mathbf{p}).$$

Relativ einfach lassen sich die obigen Erwartungswerte und Varianzen im Fall der Entropie zu einer beliebigen Basis $a > 1$ berechnen, denn aufgrund von $\log_a x = \ln x / \ln a$, $x > 0$ unterscheiden sich diese Diversitätsmaße nur um den Faktor $1/\ln a$, weshalb es genügt, den Fall der Basis $a = e$ (Eulersche Konstante bzw. natürlicher Logarithmus) zu betrachten. Wir erhalten:

$$\frac{\partial}{\partial p_i} D(\mathbf{p}) = -\ln p_i - 1, \quad \frac{\partial^2}{\partial p_i^2} D(\mathbf{p}) = -\frac{1}{p_i}$$

und Null für alle echt-gemischten partiellen Ableitungen zweiter Ordnung. Mit den obigen Formeln ergibt das

$$E(D(\hat{\mathbf{p}})) \approx D(\mathbf{p}) - \frac{(N-n)(s-1)}{2n(N-1)}$$

$$\hat{V}(\nabla D(\mathbf{p})) \approx \sum_{i=1}^s p_i \ln^2 p_i - D^2(\mathbf{p}),$$

so daß folgende Größen geeignete Schätzer für $D(\mathbf{p})$ bzw. die Varianz $\sigma^2(D(\hat{\mathbf{p}}))$ sind:

$$\hat{D}(\mathbf{p}) = D(\hat{\mathbf{p}}) + \frac{(N-n)(s-1)}{2n(N-1)}$$

$$\hat{\sigma}^2(D(\hat{\mathbf{p}})) = \frac{N-n}{n(N-1)} \left[\sum_{i=1}^s \hat{p}_i \ln^2 \hat{p}_i - D^2(\hat{\mathbf{p}}) \right]$$

mit

$$D(\hat{\mathbf{p}}) = - \sum_{i=1}^s \hat{p}_i \ln \hat{p}_i.$$

Bei der praktischen Feldarbeit ist es oft nicht möglich, den Gesamtumfang N der betrachteten Lebensgemeinschaft exakt zu ermitteln oder auch nur grob abzuschätzen. Dies ist jedoch relativ unproblematisch, wenn der Umfang n der gezogenen Stichproben klein ist gegenüber N , was in den meisten Fällen zutreffen dürfte. Bei wachsendem N und konstanten Quotienten $p_i = N_i/N$ geht die eingangs betrachtete multi-hypergeometrische Verteilung nämlich über in eine entsprechende *Multinomialverteilung*. Diese Verteilung ergibt sich auch im endlichen Fall, wenn Stichproben *mit Zurücklegen* betrachtet werden. Die zugehörigen Wahrscheinlichkeiten sind gegeben durch

$$P(X_i = n_1, \dots, X_s = n_s) = \binom{n}{n_1, \dots, n_{s-1}} p_1^{n_1} \dots p_s^{n_s}, \quad 0 \leq n_i \leq n, \quad \sum_{i=1}^s n_i = n$$

(vgl. S. 23ff). Entsprechende Rechnungen wie oben ergeben für die relativen Anteile $\hat{\mathbf{p}} = \mathbf{X}/n$:

$$E(\hat{\mathbf{p}}) = \mathbf{p}, \quad \Sigma = \frac{1}{n} \begin{pmatrix} p_1 q_1 & -p_1 p_2 & -p_1 p_3 & \dots & -p_1 p_s \\ -p_1 p_2 & p_2 q_2 & -p_2 p_3 & \dots & -p_2 p_s \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -p_1 p_s & -p_2 p_s & -p_3 p_s & \dots & p_s q_s \end{pmatrix}$$

mit $q_i = 1 - p_i$, $i = 1, \dots, s$. $\hat{\mathbf{p}}$ ist also auch hier wieder ein erwartungstreuer Schätzer für \mathbf{p} ; die Varianz-Kovarianzmatrix der Multinomialverteilung unterscheidet sich ferner nur durch einen Faktor von der Varianz-Kovarianzmatrix der multi-hypergeometrischen Verteilung. Alle obigen Rechnungen bleiben deshalb auch unter der Annahme einer Multinomialverteilung gültig, wobei lediglich der Quotient $(N-n)/(N-1)$ durch die asymptotische Konstante 1 zu ersetzen ist. Als approximative Schätzungen für $D(\mathbf{p})$ und $\sigma^2(D(\hat{\mathbf{p}}))$ ergeben sich daher folgende Größen:

$$\hat{D}(\mathbf{p}) = D(\hat{\mathbf{p}}) - \frac{1}{2n} [V(\nabla)] D(\mathbf{p})$$

$$\hat{\sigma}^2(D(\hat{\mathbf{p}})) = \frac{1}{n} \hat{V}(\nabla D(\mathbf{p})).$$

Im folgenden wollen wir uns zu diesem Themenkreis ein konkretes Beispiel ansehen. Die Daten wurden im Rahmen der *Ökosystemforschung Niedersächsisches Wattenmeer* im August 1994 erhoben und geben die gefundenen Abundanzen verschiedener adulter Arten an zwei Standorten (Versuchs- und Kontrollfläche) in einem Sandwatt wieder. Die angegebenen Schätzungen für die Diversität und die Standardabweichung $\hat{\sigma}(D(\hat{p}))$ beziehen sich auf die Entropie zur Basis e bei Multinomialverteilung.

Art	Versuch	Kontrolle
<i>Capitella</i> ¹	7	1
<i>Eteone</i> ¹	0	0
<i>Lanice</i> ¹	4	0
<i>Nephtys</i> ¹	1	1
<i>Pygospio</i> ¹	8	11
<i>Scoloplos</i> ¹	0	1
<i>Tharyx</i> ¹	1	0
<i>Crangon</i> ²	1	1
<i>Talitrus</i> ³	0	0
<i>Ostracoda</i> ⁴	10	6
Summe	32	21
$\hat{D}(\hat{p})$	1.7679	1.4908
$\hat{\sigma}(D(\hat{p}))$	0.1154	0.1955

¹ Ringelwurm

² Granat

³ Strandfloh

⁴ Muschelkrebs

Um einen Eindruck von der Variabilität der Schätzungen bei solch relativ kleinen Stichprobenumfängen zu bekommen, sind in der folgenden Tabelle die Ergebnisse einiger Simulationen dargestellt. Der tatsächliche Umfang der Lebensgemeinschaft in den untersuchten Habitaten wurde dabei fiktiv als 100 mal größer angenommen, d.h. $N = 3200$ Individuen im Versuchsfeld und $N = 2100$ Individuen in der Kontrollfläche. Die Schätzungen beziehen sich auf die multi-hypergeometrische Verteilung mit den fiktiven Abundanzen N_i hochgerechnet aus den vorigen Tabellenwerten.

Art	Sim. 1	Sim. 2	Sim. 3	Sim. 4	Sim. 5
<i>Capitella</i>	5	9	10	9	9
<i>Eteone</i>	0	0	0	0	0
<i>Lanice</i>	2	2	5	2	4
<i>Nephtys</i>	0	1	0	1	1
<i>Pygospio</i>	6	8	8	9	10
<i>Scoloplos</i>	0	0	0	0	0
<i>Tharyx</i>	2	1	2	0	2
<i>Crangon</i>	3	3	0	3	1
<i>Talitrus</i>	0	0	0	0	0
<i>Ostracoda</i>	14	8	7	8	5
Summe	32	32	32	32	32
$\hat{D}(\mathbf{p})$	1.6733	1.8009	1.6451	1.7028	1.7993
$\hat{\sigma}(D(\hat{\mathbf{p}}))$	0.1238	0.1124	0.0703	0.0988	0.1153

Simulationen für Versuchsfeld

Art	Sim. 1	Sim. 2	Sim. 3	Sim. 4	Sim. 5
<i>Capitella</i>	2	1	1	0	0
<i>Eteone</i>	0	0	0	0	0
<i>Lanice</i>	0	0	0	0	0
<i>Nephtys</i>	0	0	0	1	2
<i>Pygospio</i>	11	12	13	15	10
<i>Scoloplos</i>	2	0	0	1	0
<i>Tharyx</i>	0	0	0	0	0
<i>Crangon</i>	1	0	2	1	1
<i>Talitrus</i>	0	0	0	0	0
<i>Ostracoda</i>	5	8	5	3	8
Summe	21	21	21	21	21
$\hat{D}(\mathbf{p})$	1.4854	1.0446	1.2179	1.1655	1.3021
$\hat{\sigma}(D(\hat{\mathbf{p}}))$	0.1658	0.1154	0.1662	0.2212	0.1369

Simulationen für Kontrollfläche

Bei einem direkten Vergleich zwischen den empirischen und simulierten Daten fällt zunächst die erwartungsgemäß höhere Variabilität bezüglich der geschätzten Diversität in der Kontrollfläche auf, die in erster Linie auf den kleineren Stichprobenumfang zurückzuführen ist. Die geschätzte Standardabweichung $\hat{\sigma}(D(\hat{\mathbf{p}}))$ ist im Kontrollfeld durchschnittlich fast 50% größer als im Versuchsfeld. Mittelt man zum Vergleich mit den theoretischen Überlegungen noch die geschätzten Größen, so erhält man folgendes Bild:

Größe	Vorgabe	Mittelwert aus 5 Simulationen	Standardabweichung des Mittelwerts
$\hat{D}(\mathbf{p})$	1.7679	1.7242	0.0721
$\hat{\sigma}(D(\hat{\mathbf{p}}))$	0.1154	0.1041	0.0209

Versuchsfeld

Größe	Vorgabe	Mittelwert aus 5 Simulationen	Standardabweichung des Mittelwerts
$\hat{D}(p)$	1.4908	1.2431	0.1645
$\hat{\sigma}(D(\hat{p}))$	0.1955	0.1611	0.0379

Kontrollfläche

Wegen der relativ hohen geschätzten Standardabweichungen läßt sich bei den gegebenen Stichprobenumfängen z.B. mit der Tschebyscheff-Ungleichung (noch) keine statistisch abgesicherte Aussage über die Unterschiedlichkeit der Diversitäten an beiden Standorten machen. Legen wir aber die simulierten Stichprobenwerte zusammen, d.h. erhöhen wir den Stichprobenumfang auf das fünffache, ergibt sich folgendes Bild:

Art	Versuch	Kontrolle
<i>Capitella</i>	42	4
<i>Eteone</i>	0	0
<i>Lanice</i>	15	0
<i>Nephtys</i>	3	3
<i>Pygospio</i>	41	61
<i>Scoloplos</i>	0	3
<i>Tharyx</i>	7	0
<i>Crangon</i>	10	5
<i>Talitrus</i>	0	0
<i>Ostracoda</i>	42	29
Summe	160	105
$\hat{D}(p)$	1.6859	1.1863
$\hat{\sigma}(D(\hat{p}))$	0.0497	0.0921

Schätzung nach Simulation

Auf der Basis dieser "Daten" ließe sich ein signifikanter Unterschied zwischen den Diversitäten an den unterschiedlichen Standorten z.B. mit der Tschebyscheff-Ungleichung nachweisen. Sind nämlich X und Y zwei Zufallsgrößen mit demselben

Erwartungswert $E(X) = E(Y) = \mu$ und Varianzen $\sigma_X^2, \sigma_Y^2 > 0$, so besitzt die Differenz $Z = X - Y$ den Erwartungswert 0 und die Varianz $\sigma_Z^2 = E(Z^2) = \sigma_X^2 + \sigma_Y^2$; also folgt

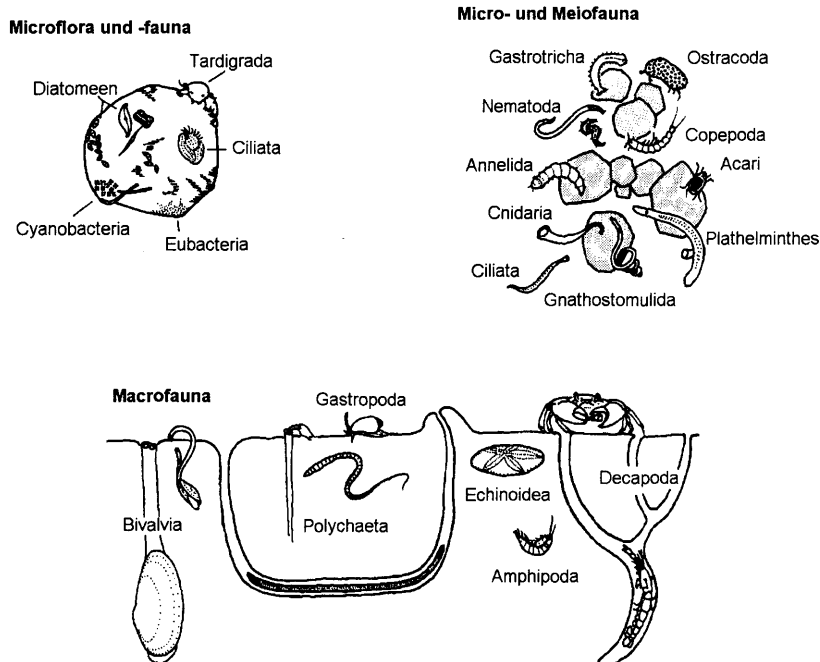
$$P(|X - Y| > c) \leq \frac{\sigma_Z^2}{c^2} = \frac{\sigma_X^2 + \sigma_Y^2}{c^2}$$

(vgl. S. 40). Zum Testen der Hypothese, ob $E(X) = E(Y)$ ist, kann also die Testgröße $|Z| = |X - Y|$ verwendet werden; die Hypothese ist – mit einer Irrtumswahrscheinlichkeit von maximal α – abzulehnen, wenn

$$|X - Y| > c_\alpha \quad \text{ist mit} \quad c_\alpha = \sqrt{\frac{\sigma_X^2 + \sigma_Y^2}{\alpha}}$$

In unserem Anwendungsbeispiel sind für X und Y die geschätzten Diversitäten und für σ_X und σ_Y die geschätzten Standardabweichungen einzusetzen. Bei $\alpha = 0.05$ erhalten wir damit $|X - Y| = 0.4996 > c_\alpha = 0.4680$; der Test lehnt also die Hypothese der Gleichheit der Diversitäten ab. Für den ursprünglichen Datensatz ergäbe dieser Test selbst für den relativ hohen Wert von $\alpha = 0.2$ lediglich $|X - Y| = 0.2771 < c_\alpha = 0.5076$, also kein Verwerfen der Hypothese.

Die folgenden, etwas umfangreicheren Datensätze zum selben Themenkreis sowie die unten dargestellten Graphiken sind der Monographie von Reise (1985), S. 14 und S. 49 entnommen. Sie geben die gefundenen Abundanzen des Zoobenthos im Sandwatt des Königshafens (Insel Sylt) zu verschiedenen Zeitpunkten wieder. Die Beprobung erfolgte mit Stechrohren über einer Probenfläche von insgesamt 10 cm^2 und einer Tiefe von 8 cm.



Art	Juni 1984	Oktober 1980
<i>Foraminifera</i> (Wurzelfüßer)	1	6
<i>Ciliata</i> (Wimpertierchen)	165	132
<i>Hydrozoa</i> (Nesseltiere)	1	0
<i>Gnathostomulida</i> (Sandwürmer)	12	2
<i>Plathelminthes</i> (Plattwürmer)	178	177
<i>Gastrotricha</i> (Rundwürmer)	18	22
<i>Nematoda</i> (Fadenwürmer)	1357	1097
<i>Nemertini</i> (Schnurwürmer)	0	4
<i>Polychaeta</i> (Röhrenwürmer)	317	60
<i>Oligochaeta</i> (Borstenwürmer)	62	5
<i>Bivalvia</i> (Muscheln)	40	5
<i>Gastropoda</i> (Schnecken)	8	0
<i>Ostracoda</i> (Muschelkrebse)	199	23
<i>Copepoda</i> (Ruderfußkrebse)	1349	257
<i>Acari</i> (Milben)	1	0
Summe	3708	1790
$\hat{D}(\mathbf{p})$	1.5706	1.3007
$\hat{\sigma}(D(\hat{\mathbf{p}}))$	0.0165	0.0274

Der obige Test zum Niveau $\alpha = 0.05$ liefert hier $|X - Y| = 0.2699 > c_\alpha = 0.1430$, d.h. die Hypothese gleicher Diversitäten wird auch hier verworfen.

Literatur

Die angegebene Literaturliste stellt lediglich eine kleine Auswahl von Titeln dar, die mit den hier behandelten Themen in engem Zusammenhang stehen, und erhebt daher keinen Anspruch auf Vollständigkeit.

- [1] M. Begon, J.L. Harper, C.R. Townsend: *Ökologie. Individuen • Populationen • Lebensgemeinschaften*. Birkhäuser, Basel 1991.
- [2] C.D. Bonham: *Measurements for Terrestrial Vegetation*. Wiley, N.Y. 1989.
- [3] N.Cressie: *Statistics for Spatial Data*. Wiley, N.Y. 1991.
- [4] P.G.N. Digby, R.A. Kempton: *Multivariate Analysis of Ecological Communities*. Population and community biology series. Chapman & Hall, London 1994.
- [5] P.J. Diggle: *Statistical Analysis of Spatial Point Patterns*. Mathematics in Biology. Ac. Press, N.Y. 1983.
- [6] S. Engen: *Stochastic Abundance Models with Emphasis on Biological Communities and Species Diversity*. Chapman & Hall, London 1978.
- [7] P. Greig-Smith: *Quantitative Plant Ecology*. Blackwell Sci. Publ., Oxford 1983.
- [8] J. Hartung, B. Elpelt, K.-H. Klösener: *Statistik. Handbuch der angewandten Statistik*. Oldenbourg Verlag, München 1989.
- [9] C.J. Krebs: *Ecology. The Experimental Analysis of Distribution and Abundance*. Harper & Row, N.Y. 1985.
- [10] U. Krenkel: *Einführung in die Wahrscheinlichkeitstheorie und Statistik*. Vieweg, Braunschweig 1991.
- [11] W. Lampert, U. Sommer: *Limnoökologie*. Thieme, Stuttgart 1993.
- [12] J. Lamprecht: *Biologische Forschung: Von der Planung bis zur Publikation*. Parey, Berlin 1992.
- [13] A.E. Magurran: *Ecological Diversity and Its Measurement*. Croom Helm, London 1988.
- [14] M. Mühlenberg: *Freilandökologie*. Quelle & Meyer, Heidelberg 1993.
- [15] G.P. Patil, C.R. Rao (Eds.): *Handbook of Statistics. Volume 12: Environmental Statistics*. North Holland, Amsterdam 1994.
- [16] E. Pielou: *Mathematical Ecology*. Wiley, N.Y. 1977.

- [17] K. Reise: *Tidal Flat Ecology. An Approach to Species Interactions*. Ecological Studies 54, Springer, Berlin 1985.
- [18] O. Richter, D. Söndgerath: *Parameter Estimation in Ecology. The Link between Data and Models*. VCH, Weinheim 1990.
- [19] G.A.F. Seber: *The Estimation of Animal Abundance and related parameters*. Griffin, London 1982.