

Richtwerte für die Datenqualität von bodenzoologischen Freiland-erhebungen

Klemens Ekschmitt

Synopsis

Data on abundances of soil fauna are collected from the literature and analysed with respect to spatial variability. In spite of a large taxonomical, ecological, and methodological multitude represented in these data, a consistent relation is found between variance and mean of the samples. Rules of thumb are derived for sample-sizes and reliability of data.

Bodentiere, Variabilität, Probenahme, Datenqualität

1. Einleitung

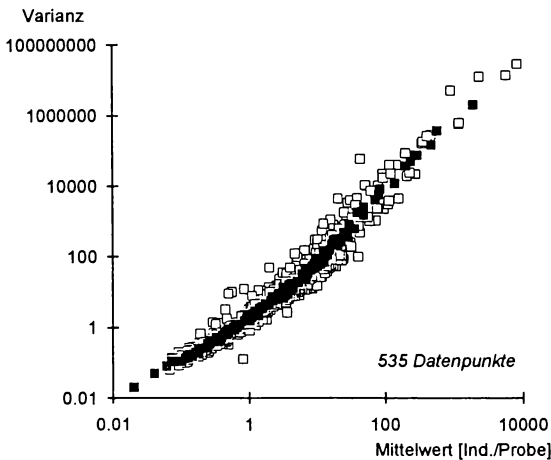
In der Literatur wird häufig die Feststellung getroffen, daß Bodentiere in der Fläche ungleichmäßig verteilt sind, und daß die erfaßten Individuenzahlen in den einzelnen Proben stark voneinander abweichen. Diese Variabilität in der räumlichen Verteilung von Bodentieren führt zu einer starken Unsicherheit bei der Versuchsplanung hinsichtlich eines angemessenen Probenaufwands und hinsichtlich der zu erwartenden Datenqualität.

Für den vorliegenden Beitrag wurden möglichst viele und umfangreiche Freiland-Erhebungen von Bodentierarten aus der Literatur zusammengetragen und die Daten in ihrer räumlichen Heterogenität untersucht. Diese Daten sollen als historische Voruntersuchungen gelten und als Erfahrungshintergrund für zukünftige Untersuchungen dienen. Im folgenden werden anhand dieser Datensammlung zwei Fragen untersucht:

- Wie stark streuen Bodentierdaten?
- Wie sicher sind die daraus gewonnenen Ergebnisse?

2. Zur Datenstreuung

In Abbildung 1 sind Varianzen und Mittelwerte aus der Datensammlung in doppelt logarithmischer Darstellung gegeneinander aufgetragen (Quellen im Anhang). Tabelle 1 gibt einen Überblick über die vertretenen Taxa.



Taxa
Nematoda
Enchytraeidae
Lumbricidae
Arachnida
Acarina
Isopoda
Diplopoda
Chilopoda
Collembola
Dermaptera
Coleoptera

Abb. 1: Zusammenhang zwischen Varianz und Mittelwert bei Bodentierarten. Jeder Datenpunkt beruht auf mindestens 30 Proben. Die schwarzen Quadrate bezeichnen das zentrale Drittel der Daten.

Tab. 1: In den Daten von Abbildung 1 repräsentierte Taxa.

Die Daten stammen aus sehr unterschiedlichen Biotopen, und die angewendeten Methoden umfassen ein breites Spektrum der in der Bodenzöologie verwendeten Probenahme- und Extraktionstechniken. Trotzdem zeigt sich ein systematischer und enger Zusammenhang zwischen Varianz und Mittelwert.

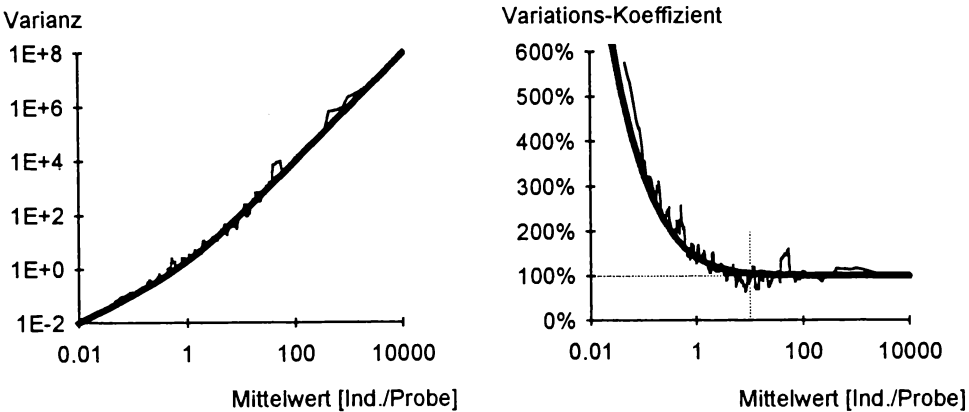


Abb. 2: Gleitende Mittelwerte aus je zehn Datenpunkten a) Varianz, b) Variationskoeffizient. Die angegebenen Formeln (s. Text) beschreiben die zentrale Tendenz der Daten mit empirisch zufriedenstellender Genauigkeit.

In Abbildung 2 sind gleitende Mittelwerte für Varianz und Variationskoeffizient über dem Mittelwert dargestellt. Die zentrale Tendenz der Abhängigkeit zwischen Datenstreuung und Mittelwert läßt sich mit zufriedenstellender Genauigkeit durch folgende Gleichungen beschreiben:

$$s^2 = \bar{x} + \bar{x}^2 \tag{1}$$

$$V = s/\bar{x} = \sqrt{1/\bar{x} + 1} \quad [100\%] \tag{2}$$

s^2 : Varianz, $\bar{x}$: Mittelwert [Ind./Probe]
V: Variationskoeffizient, s: Standardabweichung

Als geeignete Beschreibung für die Häufigkeitsverteilung von Bodentierdaten wird die negative Binomialverteilung angesehen (HARTENSTEIN 1961, BERTHET & GERARD 1965, WHEELER & al. 1987). Die Formeln stimmen mit dem Verhältnis zwischen Varianz und Mittelwert bei der negativen Binomialverteilung für k = 1 überein (z.B. PIELOU 1977, S.128).

Die Formeln entsprechen auch stochastischen Nullhypothesen für Varianz und Mittelwert bei Multinomialverteilungen. Diese gelten für eine große Klasse beliebiger Datensätze, es werden keine besonderen biologischen Zusammenhänge vorausgesetzt (HANSKI 1982, SOBERÓN & LOEVINSOHN 1987, ROUTLEDGE & SWARTZ 1991).

Ich betrachte deshalb das in Abbildung 1 und 2a ersichtliche mittlere Verhältnis zwischen Varianz und Mittelwert als unspezifisches Grundmuster, das sich aus Habitatheterogenität, Aufenthaltspräferenz und aus asynchron ablaufenden Vorgängen innerhalb der Fläche ergibt. Sofern für eine untersuchte Tiergruppe nicht spezifische Abweichungen von diesem Grundmuster bekannt sind und ein besseres Modell vorliegt, liefern die Gleichungen eine Referenz dafür, welche Varianzen im allgemeinen Durchschnitt erwartet werden können. In diesem Zusammenhang soll darauf hingewiesen werden, daß sich die Klumpungen von einzelnen Arten gegenseitig ausgleichen, wenn die Daten zusammengefaßt werden. Dadurch fällt die Streuung der Daten von Artengruppen systematisch geringer aus.

Der Variationskoeffizient, als relatives Maß für die Datenstreuung, nähert sich bei großen Mittelwerten asymptotisch dem Wert 1 (100%). Bei einem Mittelwert von 10 Individuen/Probe beträgt die Abweichung vom asymptotischen Endwert nur noch 5% (Abb. 2b). Ein Variationskoeffizient von 100% kann also als Richtwert für Daten von Bodentierarten mittlerer bis hoher Häufigkeit gelten. Die Daten von selteneren Arten streuen stärker.

Im folgenden werden die Konsequenzen dieser Datenstreuungen für die Zuverlässigkeit von Abundanzbestimmungen, für die Nachweisbarkeit von Abundanzunterschieden und für die Erfäßbarkeit von seltenen Arten dargestellt.

3. Zur Zuverlässigkeit der Ergebnisse

3.1 Erfassung von Abundanzen

Ein grundlegendes Ziel vieler Freilanduntersuchungen besteht darin, festzustellen, wie viele Tiere einer bestimmten Art sich auf der untersuchten Fläche befinden, um diese Fläche zu charakterisieren, oder als Basis für Korrelationen z.B. mit abiotischen Parametern. In Abb. 3a sind Konfidenzintervalle für 90% Sicherheit und für Mittelwerte > 10 Individuen/Probe in Abhängigkeit von der Probenzahl dargestellt.

In der Datensammlung befinden sich 17 Datensätze mit Mittelwerten über 10 Individuen/Probe, bei denen die Werte für die Einzelproben vorliegen. Aus diesen 17 Datensätzen wurden nach der bootstrap-Methode jeweils 1000 Zufallsproben mit unterschiedlichen Stichprobenumfängen gezogen, die Unterschiede zwischen dem zufällig erfaßten Mittelwert und dem Mittelwert des gesamten Datensatzes aufgelistet, und das Intervall registriert, das 90% der 1000 Fälle einschließt. Dieses Vorgehen entspricht einem empirischen Großversuch, bei dem 1000 Personen Proben mit unterschiedlichem Stichprobenumfang erheben, auswerten und die Ergebnisse miteinander vergleichen. Die in Abbildung 3a dargestellten Datenpunkte beruhen vollständig auf empirischen Daten.

Um den systematischen Zusammenhang der Datenpunkte darzustellen, ist zum Vergleich eine theoretische Kurve eingetragen, die unter Verwendung der oben dargestellten Formel und unter der (falschen!) Annahme, die Daten seien normalverteilt, berechnet wurde (s. Anhang).

Die Konfidenzintervalle fallen im einzelnen stark unterschiedlich aus. Im groben Durchschnitt wird mit einem Stichprobenumfang von 10 Proben ein Konfidenzintervall von knapp unter $\pm 60\%$ des Mittelwerts erreicht. Aus einer Untersuchung, bei der mit einem Aufwand von 10 Proben 100 Tiere pro m^2 festgestellt werden, kann also im groben Durchschnitt geschlossen werden, daß der wahre Wert für die charakterisierte Gesamtfläche mit 90% Sicherheit zwischen ca. 40 und 160 Tiere/ m^2 beträgt, bei jeder zehnten Untersuchung dieser Art liegt der wahre Mittelwert noch außerhalb dieser Spanne. Für seltene Tierarten (< 10 Ind./Probe) sind die Konfidenzintervalle entsprechend größer.

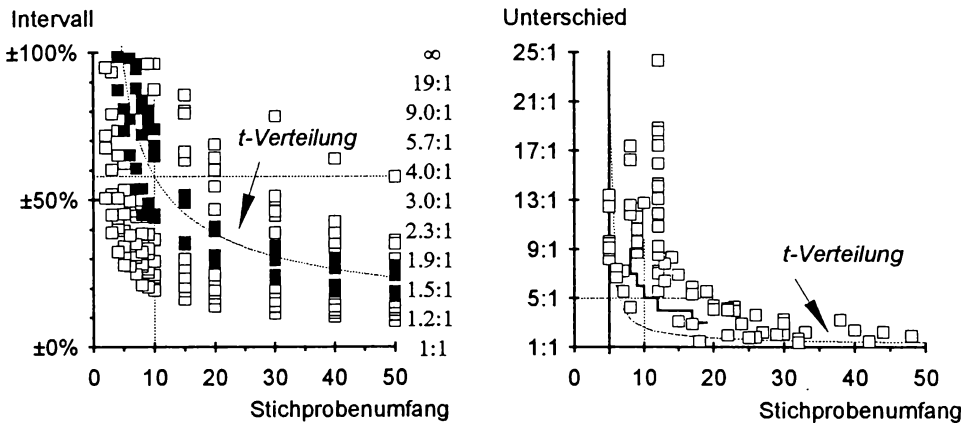


Abb. 3: Datengenauigkeit für Mittelwerte > 10 Individuen/Probe, a) Konfidenzintervalle, b) nachweisbare relative Unterschiede. Die Datenpunkte wurden aus 17 empirischen Datensätzen ermittelt, die gefüllten Quadrate bezeichnen das zentrale Drittel der Daten. Die angegebenen Kurven gelten für normalverteilte Daten (Erläuterung s. Text).

3.2 Nachweis von Abundanzunterschieden

Außer der Feststellung von Abundanzwerten, spielt der Nachweis von Abundanzunterschieden häufig eine Rolle, etwa beim Vergleich einer behandelten und einer unbehandelten Fläche. Wie groß muß ein Abundanzunterschied sein, so daß er mit einem bestimmten Probenaufwand mit einer bestimmten Sicherheit nachgewiesen werden kann?

In einem weiteren Datenexperiment wurden alle 136 möglichen Paare aus den oben erwähnten 17 Datensätzen gebildet. Für jedes Paar wurde die relative Differenz berechnet, je 1000 mal aus beiden Datensätzen Zufallsproben gezogen und t-Tests ($\alpha = 0,1$) durchgeführt. Dies wurde mit steigendem Probenumfang solange wiederholt, bis die t-Tests in 90% der 1000 Wiederholungen positiv ausfielen. Der t-Test wurde angewendet, weil er unter den gängigen Tests die größte Mächtigkeit aufweist und weil er in der Praxis häufig eingesetzt wird, obwohl die Voraussetzungen für seine Anwendung nicht zutreffen..

Dieses Datenexperiment entspricht einem empirischen Großversuch, bei dem durch massenhaft wiederholte Probenahmen festgestellt werden soll, wie häufig ein zuvor experimentell hergestellter Abundanzunterschied in den Probandaten nachgewiesen werden kann.

Die Ergebnisse sind in Abbildung 3b dargestellt. Im groben Durchschnitt läßt sich mit einem Aufwand von 2·10 Proben ein Abundanzunterschied von 5:1 mit einer Irrtumswahrscheinlichkeit von 10% ($\alpha = 0,1$) in 90% der Untersuchungen nachweisen.

Zum Vergleich wurde auch hier eine theoretische Kurve eingetragen, die unter Verwendung der oben dargestellten Formel und unter der Annahme, die Daten seien normalverteilt, berechnet wurde. Die Kurve liegt an der Untergrenze der empirischen Daten, ein deutlicher Hinweis darauf, daß die Voraussetzung für die Anwendung des t-Tests nicht erfüllt ist.

3.3 Erfassung seltener Arten

Unsicherheiten bei der Erfassung von Artenspektren ergeben sich naturgemäß aus der Erfassung der seltenen Arten. In Abbildung 4 ist dargestellt, wie häufig eine Art vorkommen muß, damit sie mit einem bestimmten Probenaufwand mit 90% Sicherheit mindestens in einem Exemplar erfaßt wird. Die Daten wurden aus der Häufigkeit von 'Nullproben' in den Datensätzen berechnet. Die durchgezogene Kurve entstammt der negativen Binomialverteilung mit $k = 1$ (s. Anhang). Sie stimmt gut mit den empirischen Daten überein.

Mit einem Aufwand von 10 Proben lassen sich noch Arten mit 90% Sicherheit erfassen, die durchschnittlich nur in jeder vierten Probe mit einem Individuum vertreten sind. Die qualitative Charakterisierung einer Fläche anhand des Artenspektrums scheint mit gleichem Probenaufwand zuverlässiger möglich zu sein, als auf der Basis von quantitativen Messungen der Abundanzen.

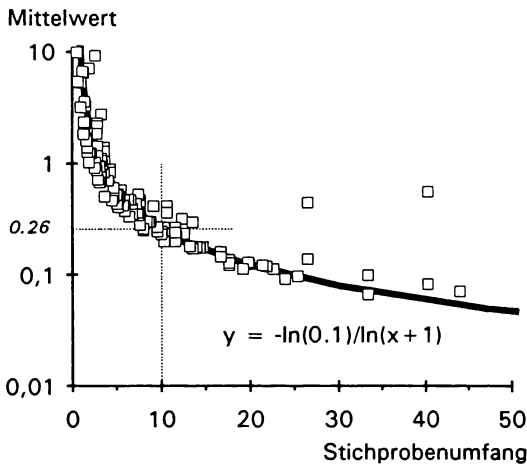


Abb. 4: Erfäßbarkeit von seltenen Arten. Die Datenpunkte wurden aus empirischen Datensätzen berechnet. Die Kurve gilt für die negative Binomialverteilung mit $k = 1$. Die Übereinstimmung ist sehr gut.

4. Diskussion

Die vorliegende Datenanalyse zeigt, daß bodenzoologische Daten einem allgemeinen Grundmuster folgen, für das spezielle statistische Methoden evaluiert bzw. entwickelt werden können. Von der Anwendung besser angepaßter Auswertungsmethoden können graduelle Verbesserungen für die Einschätzung der Datenzuverlässigkeit und für die Nachweisbarkeit von Unterschieden erwartet werden, sie werden jedoch nicht zu neuen Größenordnungen der Datenqualität hinführen.

Mit einem üblichem Probenaufwand können im allgemeinen nur sehr 'weiche' Daten über Bodentierabundanz im Freiland gewonnen werden. Typische Konfidenzintervalle von $\pm 50\%$ ($\alpha = 0,1$) liegen weit jenseits dessen, was für Labordaten und für technische Angaben als akzeptabel gilt. Die hohe Datenunsicherheit schränkt die Interpretierbarkeit von Abundanzdaten stark ein. Sie muß Berücksichtigung finden, wenn solche Daten in Modellen weiterverarbeitet werden. Dieser Aspekt wird von POETHKE & al. (1993, dieser Band) erörtert.

Literatur

- BERTHET, P. L. & G. GERARD, 1965: A statistical study of microdistribution of Oribatei (Acari). - Part I. The distribution pattern. - *Oikos* 16: 214-27.
- HANSKI, I., 1982: On patterns of temporal and spatial variation in animal populations. - *Ann. Zool. Fenn.* 19: 21-37.
- HARTENSTEIN, R., 1961: On the distribution of forest soil micro-arthropods and their fit to 'contagious' distribution functions. - *Ecology* 42: 190-4.
- PIELOU, E. C., 1977: *Mathematical Ecology*. - Wiley, New York.
- POETHKE, H.J., OERTL, D & A. SEITZ, 1993: Variabilität in ökologischen Systemen. Konsequenzen für den Zeithorizont von Modellprognosen. - *Verh. Ges. Ökol.* 22: 457-464.
- ROUTLEDGE, R. D. & T. B. SWARTZ, 1991: Taylor's Power Law re-examined. - *Oikos* 60: 107-112.
- SOBERÓN M. J. & M. LOEVINSON, 1987: Patterns of variations in the numbers of animal populations and the biological foundations of Taylor's law of the mean. - *Oikos* 48: 249-252.
- TAYLOR, L. R., 1986: Synoptic Dynamics, Migration and the Rothamsted Insect Survey. Presidential address to the British Ecological Society, December 1984. - *Journal of Animal Ecology* 55: 1-38.
- WHEELER, T. A., KENERLEY, C. M., JEGER, M. J. & J. L. STARR, 1987: Effect of quadrat and core sizes on determining the spatial pattern of *Criconebella sphaerocephalus*. - *Journal of Nematology* 19: 413-419.

Anhang

A: Quellen-Angaben zu den Daten

- ABRAHAMSEN, G., 1969: *Oikos* 20: 54-66.
- ABRAHAMSEN, G., 1970: *Oikos* 21: 276-284.
- ALBERT, A. M., 1982: *Pedobiologia* 23: 337-347.
- ALBERT, R., 1982: Hochschulsammlung Naturwissenschaft: Biologie Bd. 16. - Hochschulverlag, Freiburg (Brsgr.)
- BERBIERS, Ph., MAELFAIT, J.-P. & J. MERTENS, 1989: *Rev. Écol. Biol. Sol.* 26: 305-320.
- BERTHET, P. & G. GERARD, 1965: *Oikos* 16: 214-227.
- CHALUPSKY, Jr. J. & J. LEPS, 1985: *Oecologia* 68: 153-157.
- COLE, L. C., 1946: *Ecol. Monogr.* 16: 49-86.
- DAVID, J. F. & T. COURET, 1985: *Rev. Écol. Biol. Sol* 22: 367-380.
- DAVIS, R. C & S. L. SUTTON, 1977: *Ecol. Bull. (Stockholm)* 25: 45-55.
- DEBAUCHE, H. R., 1958: *Quest. Sci.* 129: 58-89.
- DESENDER, K., 1988: *Rev. Écol. Biol. Sol* 25: 101-113.
- FRANCL, L. J., 1986: *Journal of Nematology* 18: 183-189.
- GOODELL, P. & H. FERRIS, 1980: *J. Nematol.* 12: 136-141.
- GUILD, W. J. McL., 1952: *J. Anim. Ecol.* 21: 169-181.
- HAIRSTON, N. G., HILL, R. W. & U. RITTE, 1971: *Statistical Ecology* 1: 337-356.
- HARTENSTEIN, R., 1961: *Ecology* 42: 190-194.
- HARTMANN, P., 1979: Dissertation Univ. Göttingen.
- HENGVELD, R., 1979: In: R. M. CORMACK & J. ORD (eds.): *Spatial and Temporal Analysis in Ecology*. - International Cooperative Publishing House, Fairland, Maryland: 333-346.
- JUDAS, M., 1989: *Pedobiologia* 33: 339-354.
- KOEHLER, H., 1989 - unpubl. Daten, Univ. Bremen.
- MCGROARTY, D. L. & B. A. CROFT, 1978: *The Canadian Entomologist* 110: 785-794.
- NOE, J. P. & C. L. CAMPBELL, 1985: *Journal of Nematology* 17: 86-93.
- OVERMARS, H., 1985: Horizontale en verticale verdeling van nematoden in de bodem, met het oog op hun bemonsstering. - unpubl. Bericht, Landbouw-Universiteit Wageningen (Niederlande).
- REISE, K. & WEIDEMANN, G., 1975: *Pedobiologia* 15: 106-128.
- SATCHELL, J. E., 1967: In: BURGESS, A. & F. RAW (eds.): *Soil Biology*: 259-322, Academic Press, London.
- TAKEDA, H., 1979: *Res. Popul. Ecol.* 21: 120-134.
- TOPP, W., 1971: *Ann. Zool. Fennici* 8: 194-222.

VERHOEF, H. A. & A. J. VAN SELM, 1983: Holarctic Ecology 6: 387-394.
 WHEELER, T. A., KENERLEY, C. M., JEGER, M. J. & J. L. STARR, 1987: Journal of Nematology 19: 413-419.
 WORTMANN, A., 1990: Die Collembolenfauna eines biologisch dekontaminierten Bodens: Untersuchungen zur repräsentativen Erfassung und zur Indikatoreignung. - unveröff. Diplomarbeit, Univ. Bremen.
 ZAGORODNEVA, D. S. & L. L. KUZ'MIN, 1978: Pedobiologia 18: 180-188 (russisch).

B: Berechnungen

a) Konfidenzintervalle

Die Schätzungleichung für normalverteilte Daten lautet :

$$\bar{x} - t_{\nu, \alpha} \cdot s / \sqrt{n} \leq \mu \leq \bar{x} + t_{\nu, \alpha} \cdot s / \sqrt{n} \Leftrightarrow \frac{|\bar{x} - \mu|}{\bar{x}} \leq t_{\nu, \alpha} \cdot \frac{s}{\bar{x} \sqrt{n}}$$

Einsetzen von Formel 1 (s. Text):

$$s^2 = \bar{x} + \bar{x}^2 \Leftrightarrow s = \sqrt{\bar{x} + \bar{x}^2} \Rightarrow K = \frac{|\bar{x} - \mu|}{\bar{x}} = t_{\nu, \alpha} \cdot \frac{\sqrt{\bar{x} + \bar{x}^2}}{\bar{x} \cdot \sqrt{n}} = t_{\nu, \alpha} \cdot \sqrt{\frac{1/\bar{x} + 1}{n}}$$

b) t - Test

Die Prüfungsungleichung für den t - Test lautet: $|\bar{x}_2 - \bar{x}_1| / \sqrt{\frac{s_1^2 + s_2^2}{n}} \geq t_{\nu, \alpha}$

Die Näherung für große Mittelwerte ergibt: $s = \sqrt{\bar{x} + \bar{x}^2} \Rightarrow s \approx \bar{x}; \quad \bar{x} > 10$

Einsetzen in die Prüfungsungleichung und dividieren durch $\bar{x}_1$:

$$\frac{|\bar{x}_2 - \bar{x}_1|}{\sqrt{\frac{\bar{x}_1^2 + \bar{x}_2^2}{n}}} \geq t_{\nu, \alpha} \Leftrightarrow \frac{|\bar{x}_2 / \bar{x}_1 - 1|}{\sqrt{\frac{\bar{x}_2^2 / \bar{x}_1^2 + 1}{n}}} \geq t_{\nu, \alpha} \Leftrightarrow \frac{|D - 1|}{\sqrt{\frac{D + 1}{n}}} \geq t_{\nu, \alpha}; \quad D = \bar{x}_2 / \bar{x}_1$$

Quadrieren und Lösen nach D ergibt: $D = \frac{n + \sqrt{2nt_{\nu, \alpha}^2 - t_{\nu, \alpha}^4}}{n - t_{\nu, \alpha}^2}; \quad \nu = n - 1; \quad \alpha: 0.1$

c) Seltene Arten

Die empirische Erfassungswahrscheinlichkeit errechnet sich aus der Häufigkeit von

"Nullproben" in den Datensätzen: $P(x > 0) = 1 - h(0)^n$

Die theoretische Wahrscheinlichkeit wurde nach der negativen Binomialverteilung berechnet :

$$P(y) = \binom{k+y-1}{y} \cdot p^k \cdot q^y; \quad q = 1 - p; \quad \mu = k \cdot q / p; \quad p = \frac{k}{\mu + k}$$

Einsetzen von $k = 1$ und $y = 0$ (Nullproben) ergibt:

$$P(0) = \binom{1+0-1}{0} \cdot p^1 \cdot q^0 \Leftrightarrow P(0) = p; \quad P(0) = \frac{1}{\mu + 1}$$

Die Wahrscheinlichkeit α , in n Proben kein einziges Individuum zu erfassen beträgt :

$$\alpha = P_n(0) = (\mu + 1)^{-n} \Leftrightarrow \ln(\alpha) = -n \cdot \ln(\mu + 1) \Leftrightarrow n = -\ln(\alpha) / \ln(\mu + 1)$$

Adresse

Klemens Ekschmitt, Universität Bremen, FB 2 (Biologie) AG Ökosystemforschung und Bodenökologie, Leobenerstraße, D-W-2800 Bremen 33

ZOBODAT - www.zobodat.at

Zoologisch-Botanische Datenbank/Zoological-Botanical Database

Digitale Literatur/Digital Literature

Zeitschrift/Journal: [Verhandlungen der Gesellschaft für Ökologie](#)

Jahr/Year: 1993

Band/Volume: [22_1993](#)

Autor(en)/Author(s): Ekschmitt Klemens

Artikel/Article: [Richtwerte für die Datenqualität von bodenzoologischen Freilandhebungen 451-456](#)