



Statistické hodnocení biodiverzity

Jiří Jarkovský, Danka Haruštiaková

Institut biostatistiky a analýz, Masarykova univerzita, 2014



Program kurzu

- 25.9 Biodiverzita jako pojem + Biodiverzita a biostatistika
- 2.10. Vizualizace biodiverzity
- 9.10. Indexy diverzity a jejich statistická spolehlivost
- 16.10. Species-abundance křivky a stochastické modely
- 23.10. Niche-oriented species - abundance modely
- 30.10. Aplikovatelnost parametrických a neparametrických statistických technik při hodnocení biodiversity
- 6.11. Aplikovatelnost parametrických a neparametrických statistických technik při hodnocení biodiversity
- 13.11. Možnosti frakcionace biologických společenstev a následná analýza biodiversity získaných podjednotek
- 20.11. Parametrické hodnocení biodiversity ve vícerozměrných analýzách
- 27.11. Parametrické hodnocení biodiversity ve vícerozměrných analýzách
- 4.12. Případová studie: Parazitární společenstva
- 11.12. Případové studie: Lišejníky a znečištění ovzduší
- 18.12. předtermín



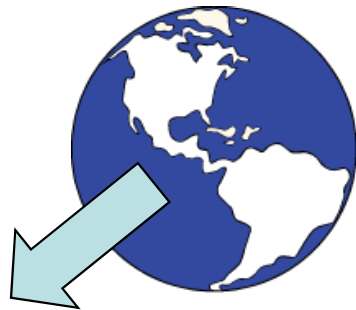


V-VI. Stochastické a na niku orientované species abundance models

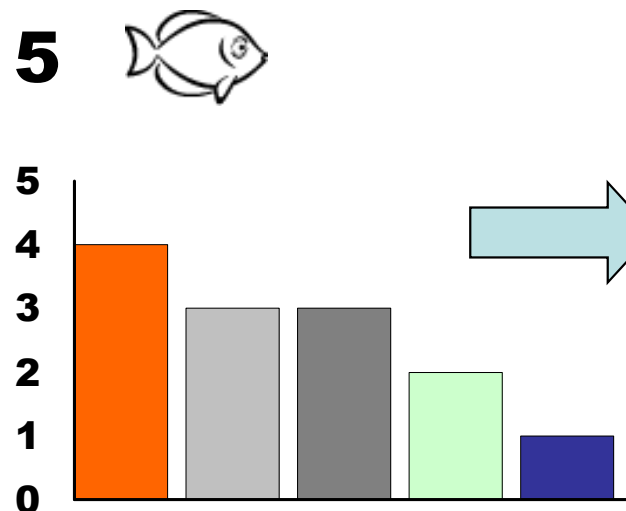
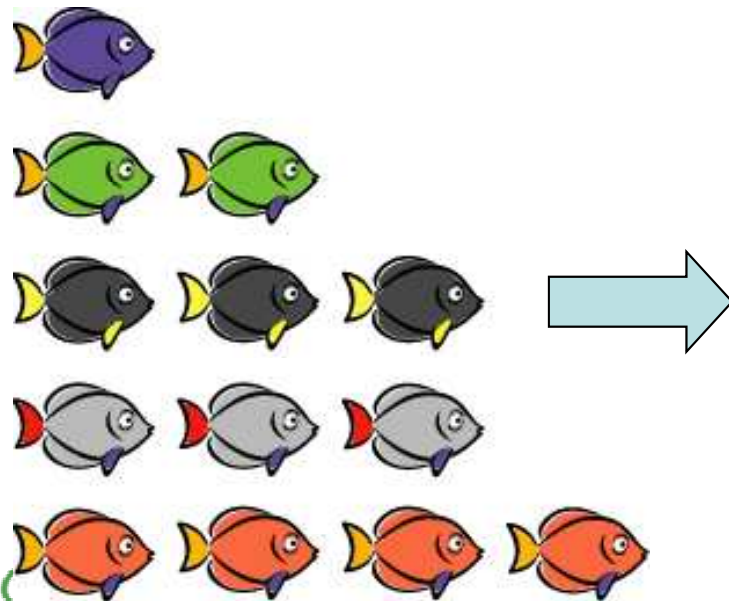


Biodiverzita ?

- ◆ Místo na Zemi → žijí zde organismy, tj. je zde biodiverzita → jak ji popsat, vysvětlit a co to znamená ?



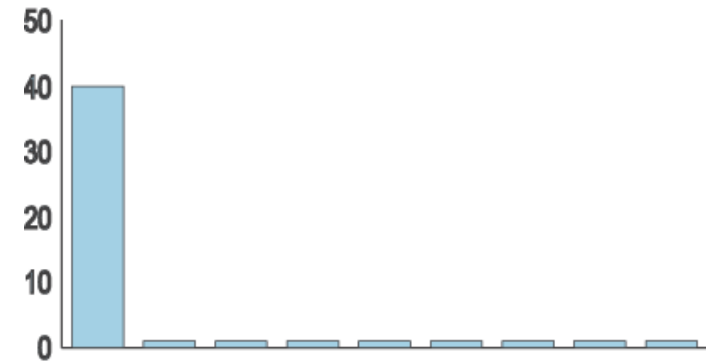
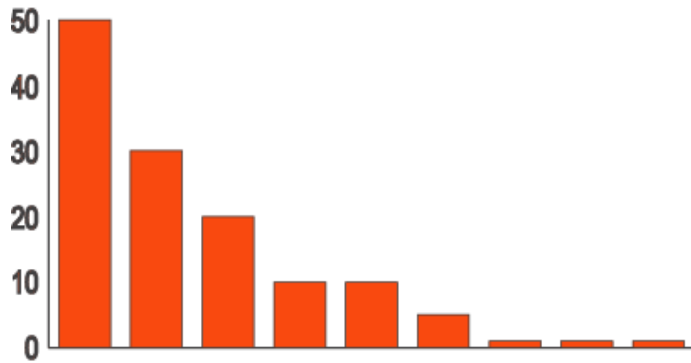
- Dvě složky biodiverzity:
 - I. Různorodost – počet různých organismů (kvalita)
 - II. Relativní abundance – poměr výskytu organismů (kvantita)



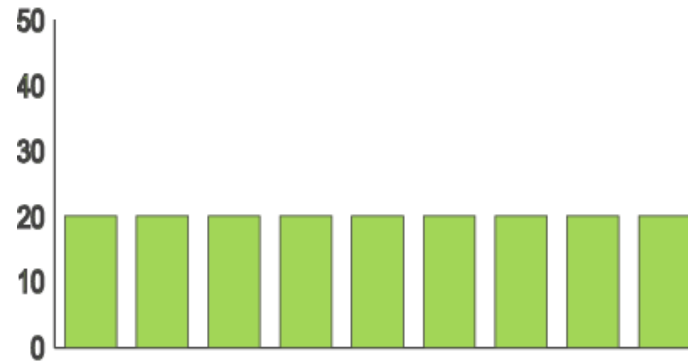
Proč ?



Species abundance curves

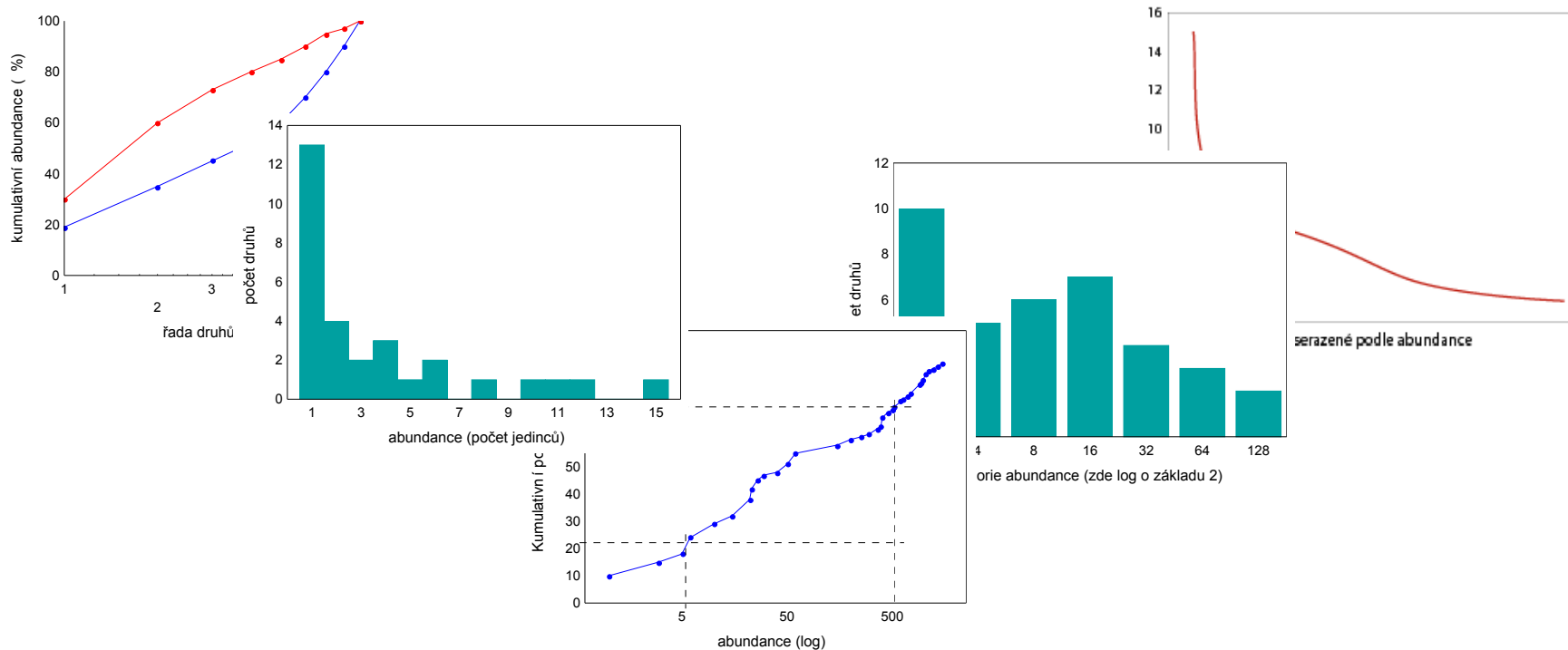


?



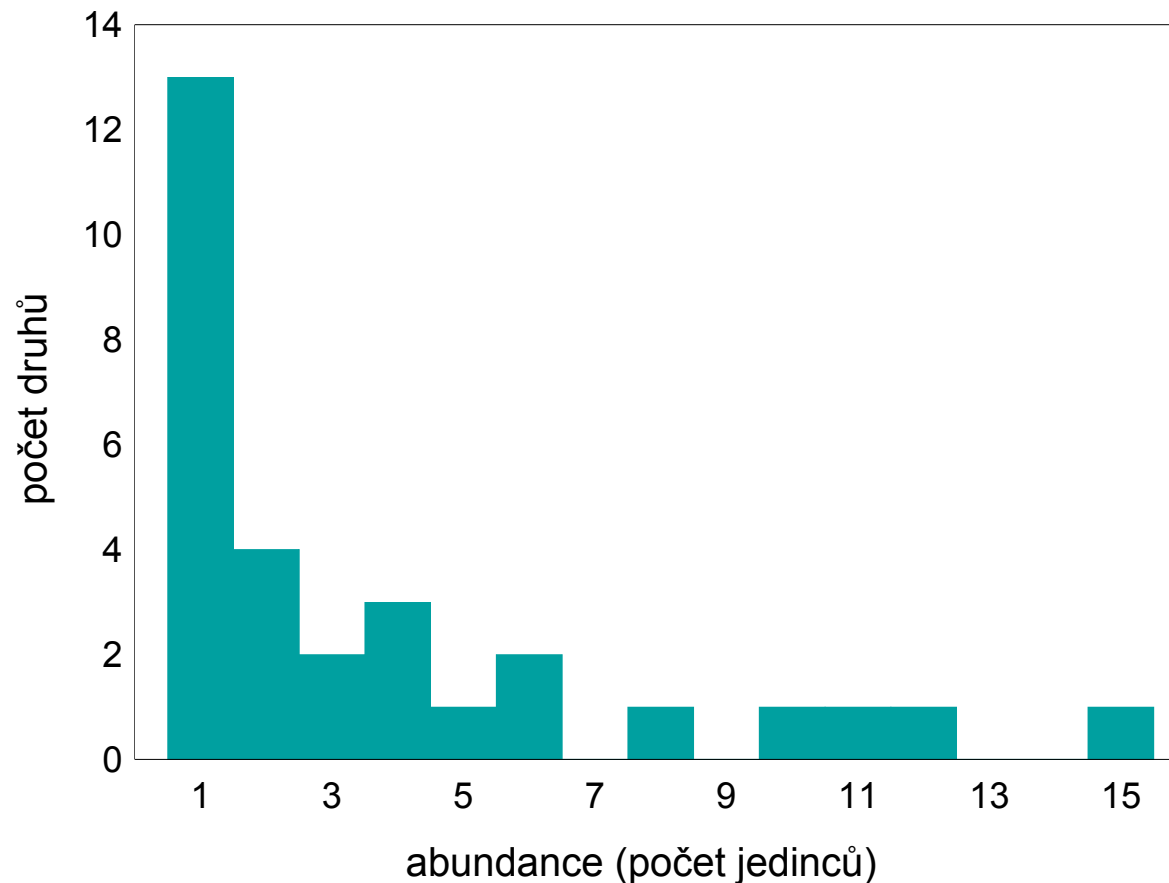
Zobrazení křivky společenstva

- Různé metody zobrazení – různé pohledy na společenstvo
- Jednoduché optické srovnání různých společenstev nebo modelových průběhů společenstev
- Jednoduchá tvorba v Excelu



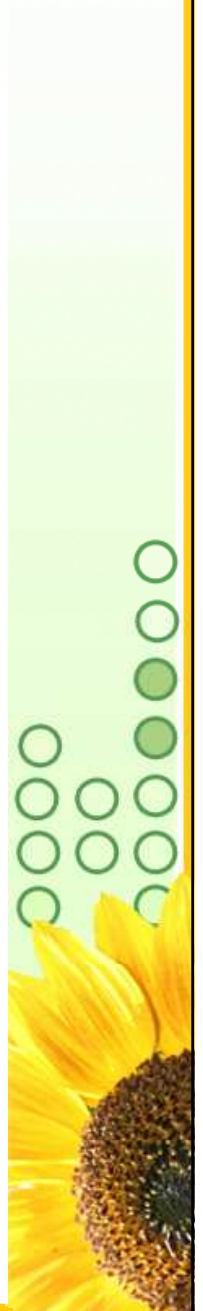
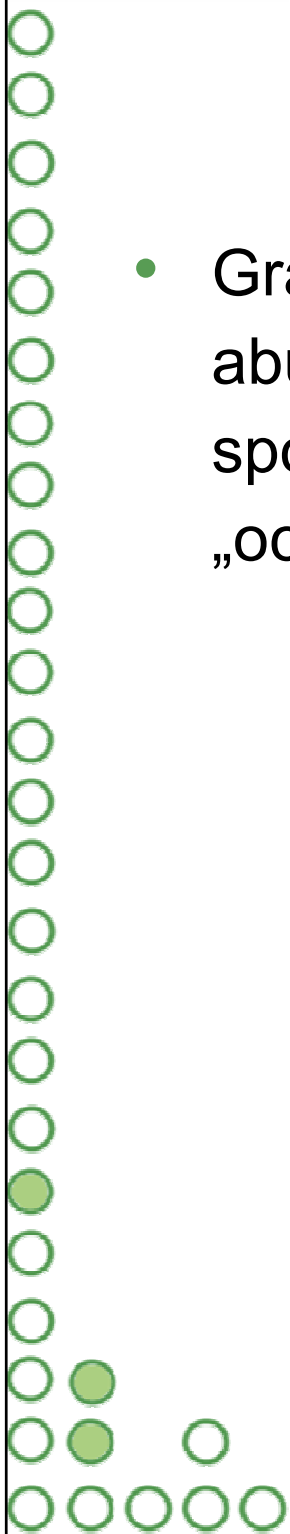
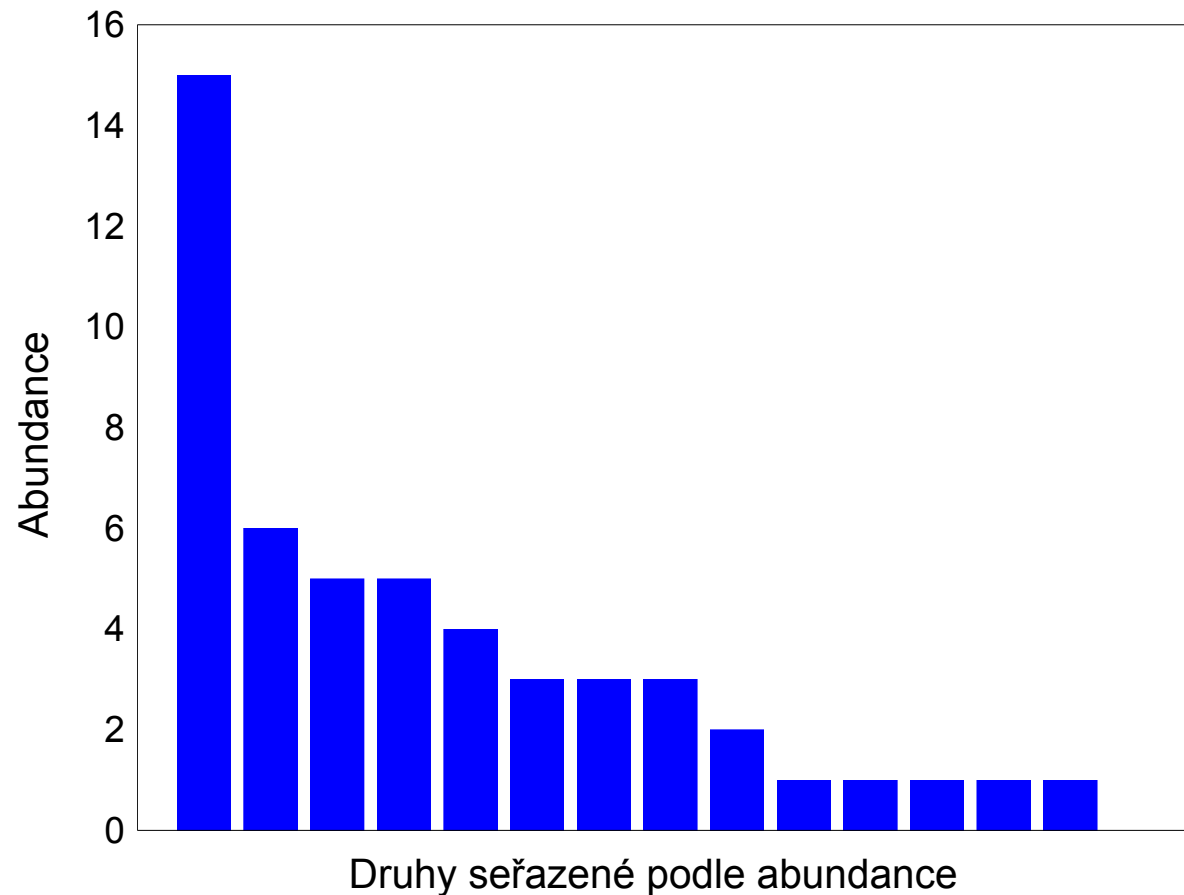
Četnost druhů

- Četnost druhů s danou abundancí je vynášena proti hodnotám nalezených abundancí. Graf poskytuje uživateli přehled o rozložení vzácných, středně početných a hojně se vyskytujících druhů.



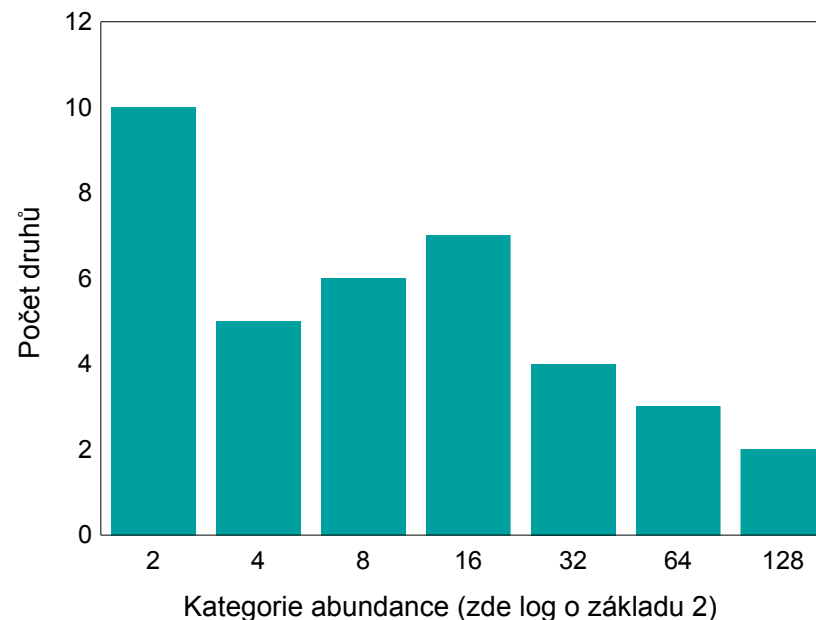
Rank abundance plot

- Graf zobrazuje abundanci druhů seřazených podle této abundance. Poskytuje uživateli přehled o tvaru společenstva – vyrovnanost abundancí, přítomnost „ocásku“ vzácných druhů apod.



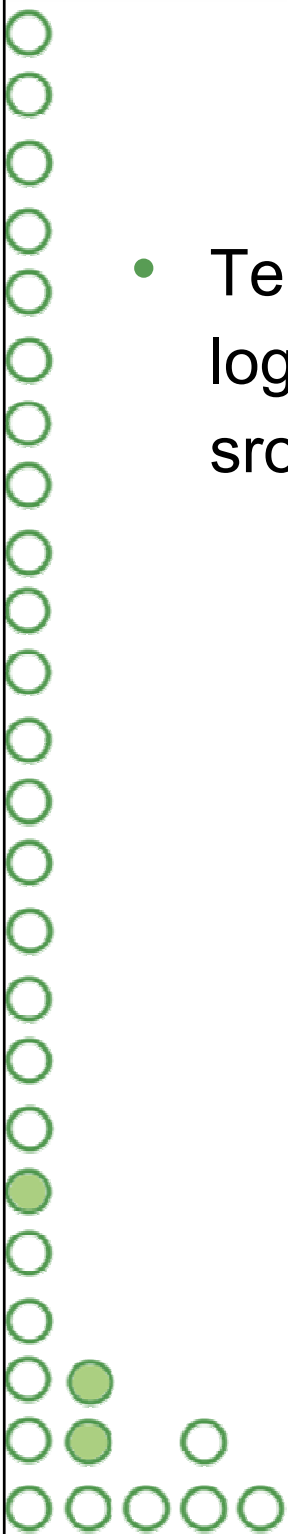
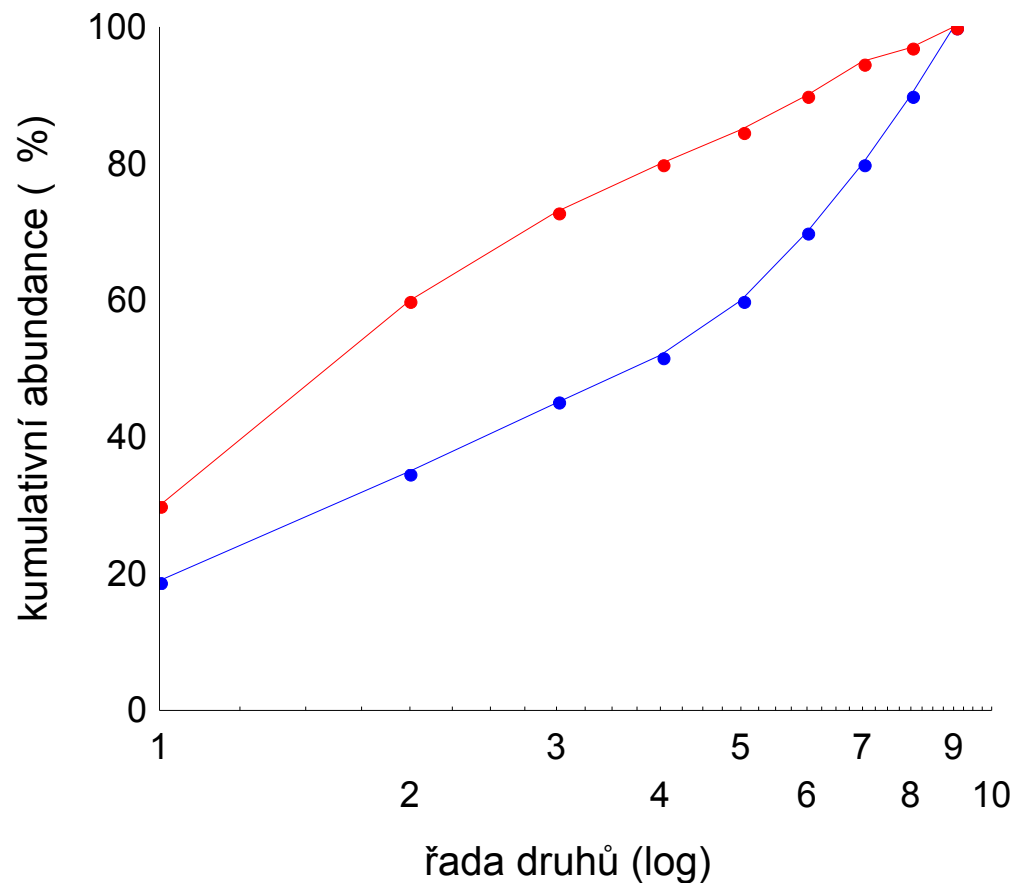
Graf kategorií abundance

- Je obdobou grafu četnosti druhů s určitou abundancí, namísto konkrétních hodnot zde jsou ale třídy abundance a četnost druhů k nim náležejících. Opět umožňuje sledovat relativní podíly vzácných a hojných druhů. Ve formě početnosti druhů v log třídách abundance jsou generovány výsledky některých species abundance models.



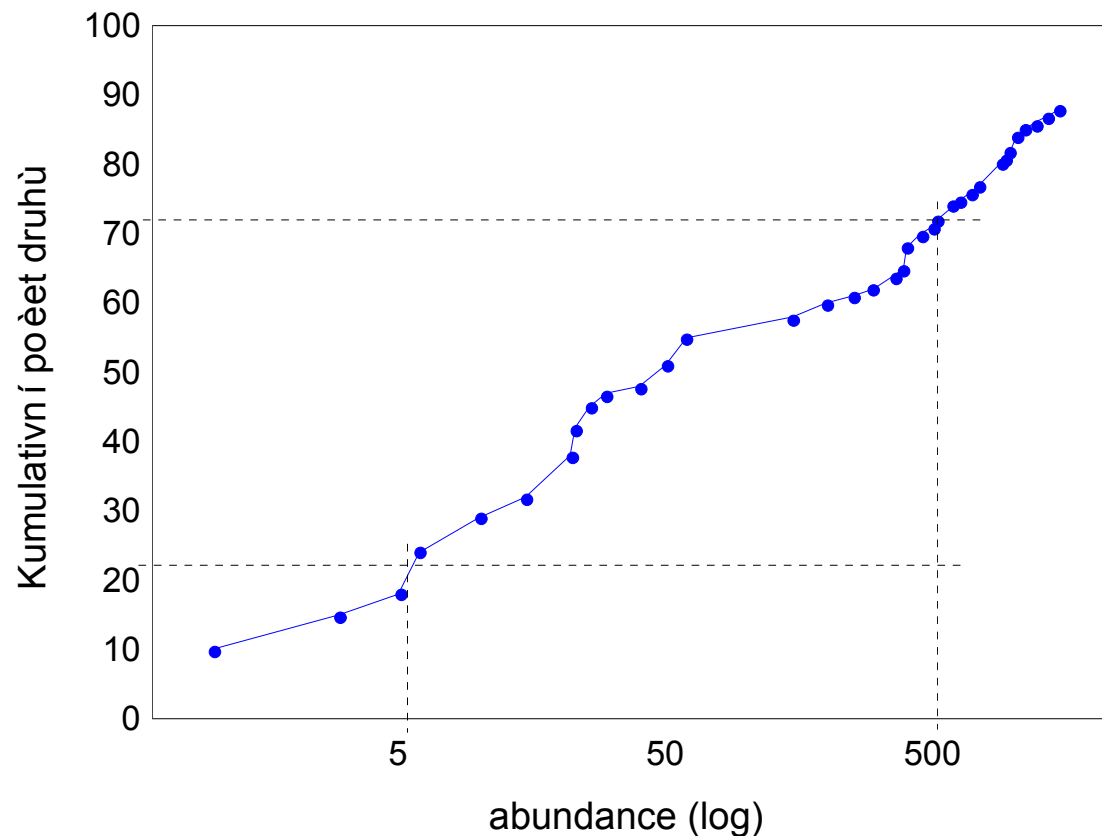
K- dominance plot

- Tento graf vynáší kumulativní abundanci druhů proti logaritmu druhové řady. Může být využit pro optické srovnání diverzity různých vzorků



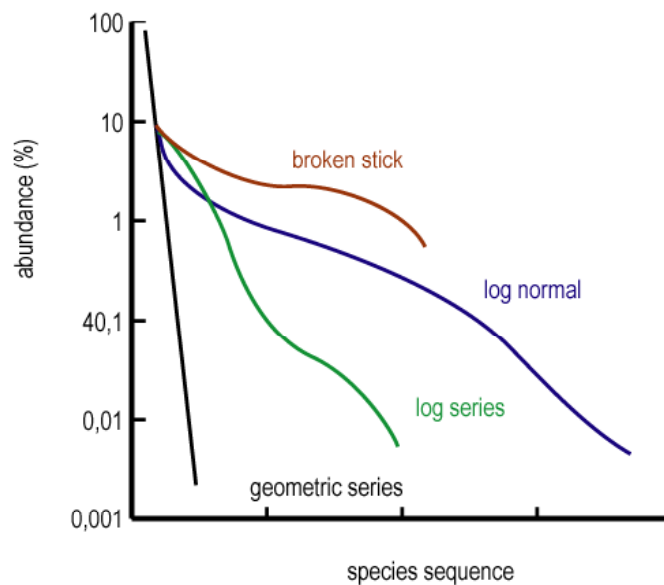
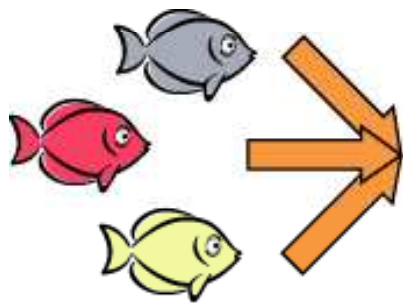
Kumulativní počet druhů

- Graf vynáší kumulativní počet druhů proti ose jejich logaritmované abundance. Graf slouží jako doplněk k výpočtu Q statistiky, který je založen na obdobně uspořádaných datech. Zobrazuje strmost narůstání počtu druhů se stoupající abundancí.



Species abundance models

- Modely, jak by měly být abundance druhů teoreticky rozloženy
- Rozložení abundancí modelu odráží **ekologické předpoklady modelu**
- Porovnání reálných abundancí s teoretickým rozložením má zjistit zda reálné společenstvo odpovídá některému z teoretických modelů, tj. způsobů vytváření společenstev (**ekologické procesy**)



?

=



Species abundance models

Matematické modely

- velká přesně nedefinovaná společenstva
- matematicky definovaná rozložení abundancí
- jednoduše testovatelné

Biologické modely

Orientované na niku

- malá společenstva
- obtížné testování
- není přesné matematické vyjádření (pouze některé)

Ostatní

- např. modely založené na rychlosti kolonizace, rozmnožování a úhynu organismů

Některé modely mohou zároveň patřit do více skupin (např. mohou být na niku orientované a mít přesné matematické vyjádření) nebo některé modely popisují různými způsoby stejný výsledek
Základ modelu není jeho zařazení do skupiny, ale co by měl reálně znamenat

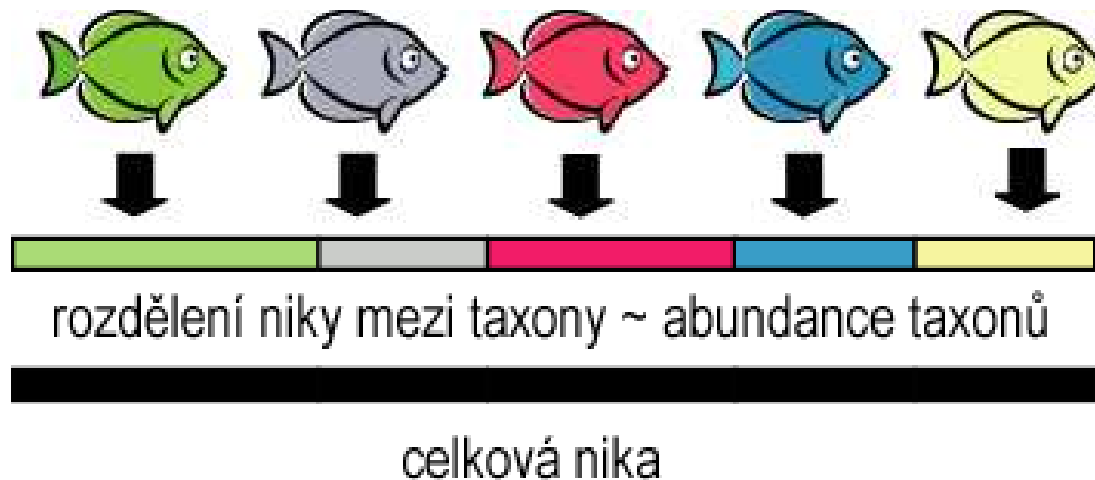
Problémem některých modelů je, že si jsme schopni je představit pouze teoreticky, ne už výpočet.



Předpoklady modelů

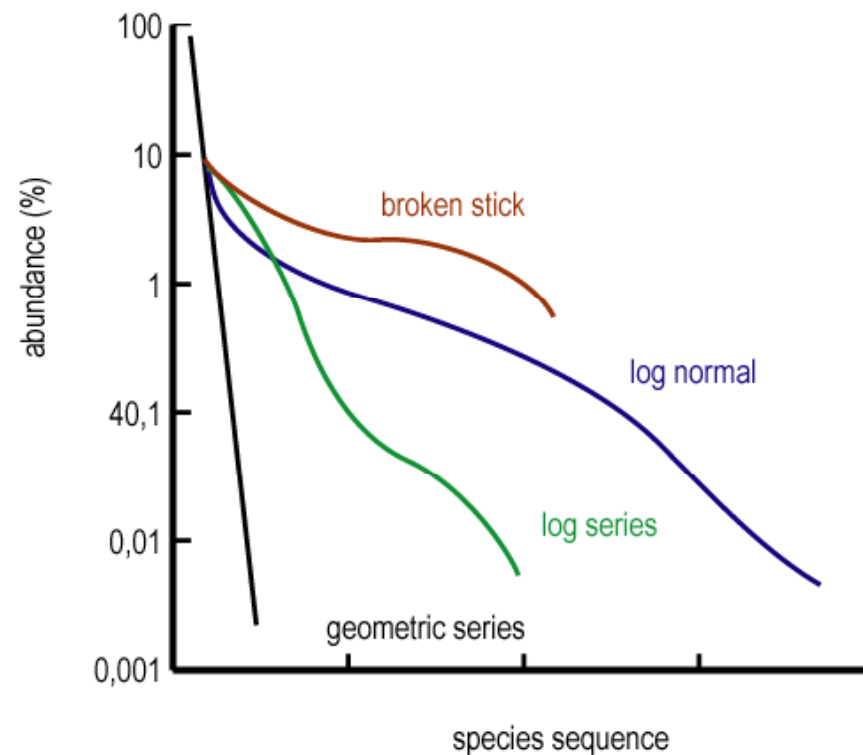
Na lokalitě existuje určitý životní prostor – nika

- zde žijící organismy si tento životní prostor dělí → abundance taxonů teoreticky odpovídá jejich podílu, který si na lokalitě získaly
- K dělení dochází na základě ekologických vlastností organismů (nároky na podmínky prostředí, životní strategie, kolonizace, kompetice atd.), tj. každý model odráží jinou teoretickou situaci (vlastnosti prostředí a taxonů) podílející se na výsledné kombinaci abundancí taxonů
- Problémem modelů je, že odráží naši jednoduchou představu o probíhajících procesech – skutečnost může být mnohem složitější



Matematické modely

- Přesné matematické vyjádření
- Geometrická serie, logaritmická série, log-normální rozložení, broken stick
- V tomto pořadí klesá dominance ve společenstvu a stoupá vyrovnanost
- Některé mají zároveň obdobu i ve stochastických na niku orientovaných modelech (lze je vyjádřit deterministicky i stochasticky)

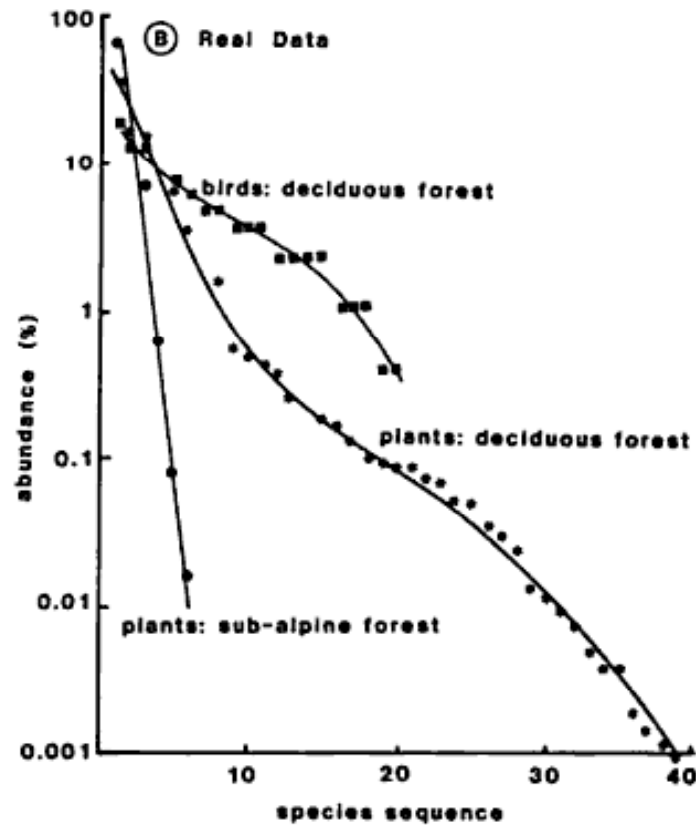
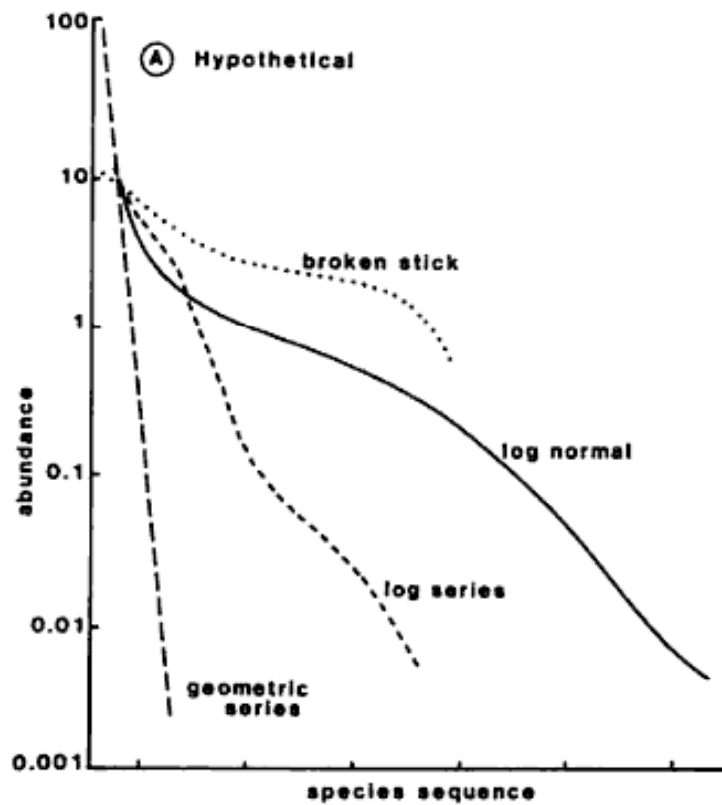


Matematické modely

Rank abundance plots – typické tvary čtyř modelů

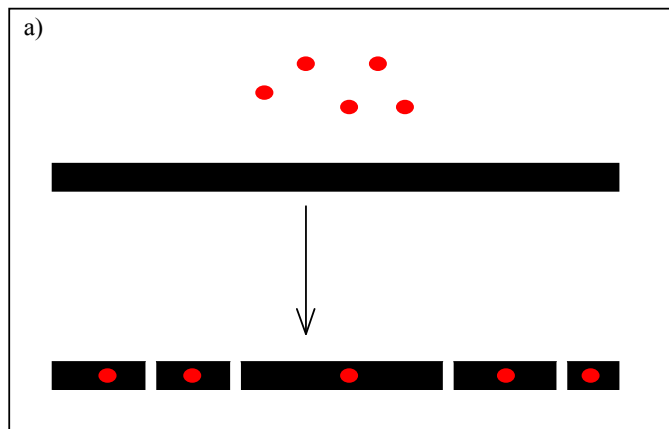
A: Hypotetické křivky

B: Tři příklady reálných společenstev

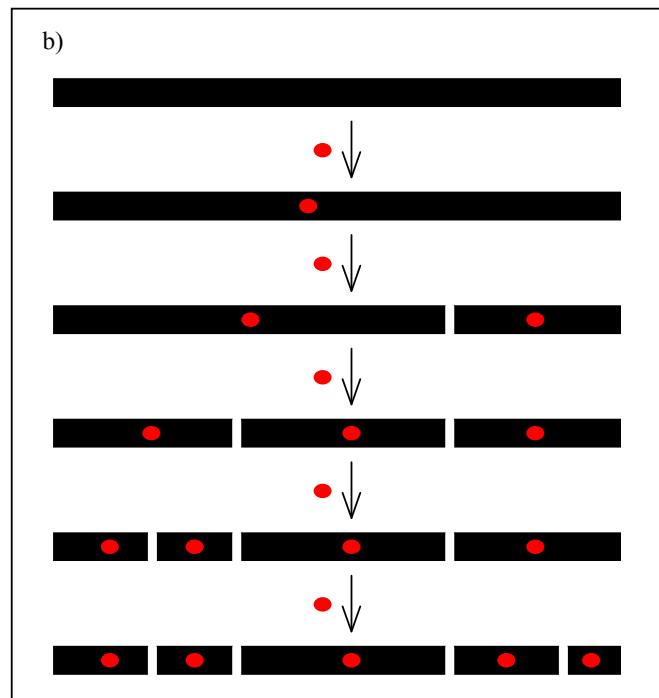


Na niku orientované modely

- Část z nich též tzv. Tokeshiho modely
- Myšlenkové konstrukce jak může být nika dělena
- Výpočet založen na stochastickém modelování (generování náhodných společenstev podle podmínek modelu)
- Různé možné způsoby dělení niky (tyto modely využívají postupné dělení, protože je jednodušší postavit modely na něm)



Simultánní dělení niky

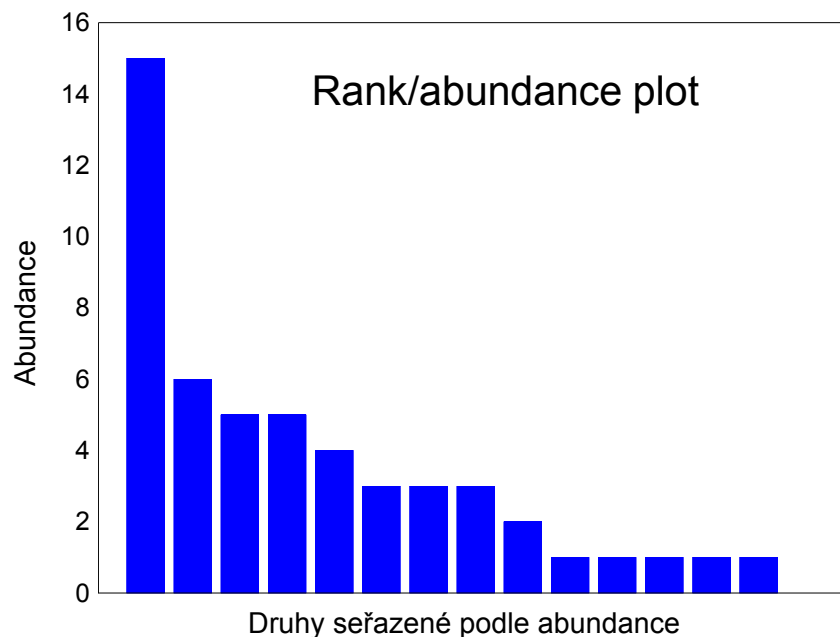


Postupné dělení niky



Geometric series

- Geometrické řady se nacházejí u druhově chudých společenstev nebo u společenstev v ranném stadiu sukcese a je založena na předpokladu, že nejdominantnější druh obsadí určitou poměrnou část zdrojů, druhý stejnou poměrnou část zbytku a tak dále
- první druh obsadí část k ($k \in (0,5;1,0)$) celkové niky, druhý druh část k' zbytku, třetí druh část k'' toho, co zůstane po umístění prvního a druhého druhu, atd.



$$i\text{-tý druh: } n_i = N C_k k (1-k)^{i-1}$$

$$C_k = [1 - (1-k)^s]^{-1}$$

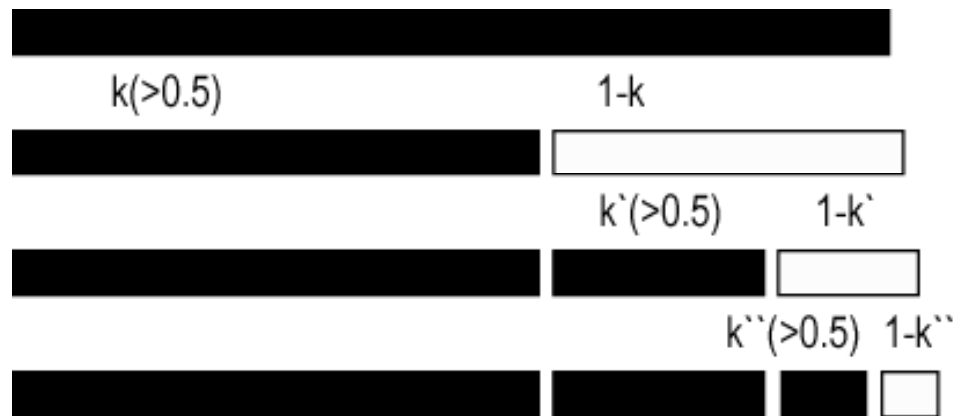
Konstanta zajišťující $\sum n_i = N$

$$k : \frac{N_{\min}}{N} = \frac{k(1-k)^s}{1 - (1-k)^s}$$



Geometric series

- Geometrické řady se nacházejí u druhově chudých společenstev nebo u společenstev v ranném stadiu sukcese a je založena na předpokladu, že nejdominantnější druh obsadí určitou poměrnou část zdrojů, druhý stejnou poměrnou část zbytku a tak dále
- první druh obsadí část k ($k \in (0,5; 1,0)$) celkové niky, druhý druh část k' zbytku, třetí druh část k'' toho, co zbude po umístění prvního a druhého druhu, atd.



$$i\text{-tý druh: } n_i = N C_k k (1-k)^{i-1}$$

$$k : \frac{N_{\min}}{N} = \frac{k(1-k)^S}{1-(1-k)^S}$$



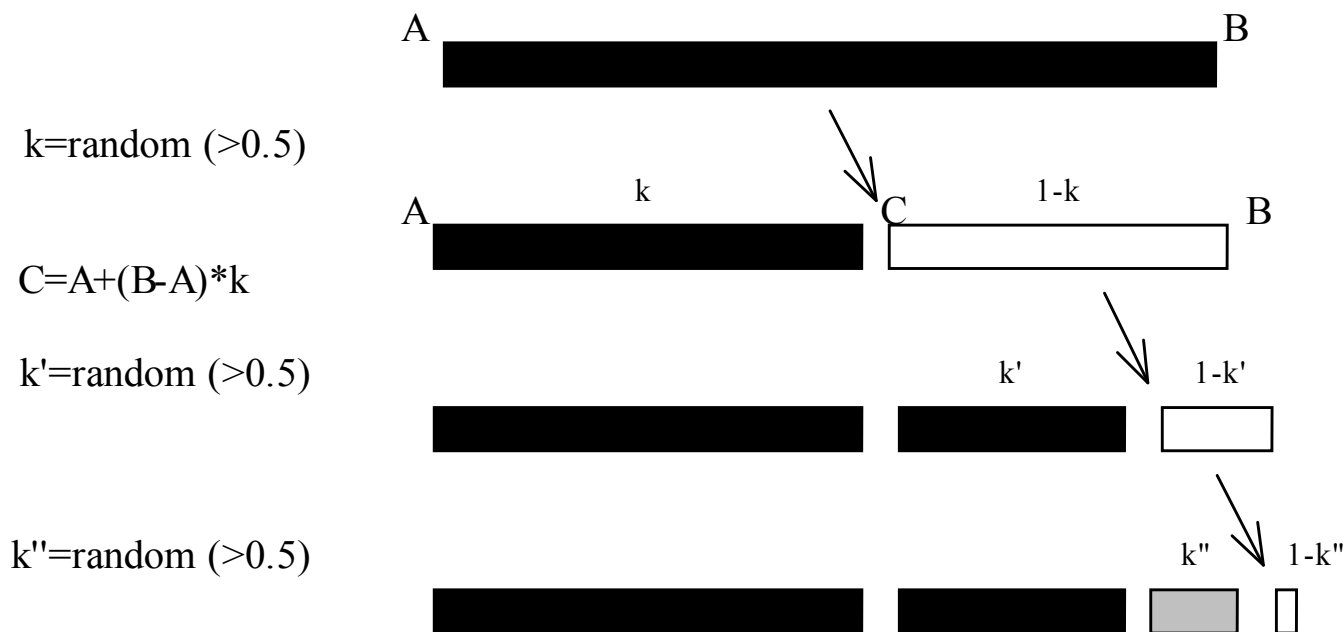
Geometric series

- Tento model byl vypracován jako první z modelů dělení niky. Jeho autorem je Motomura (1932). Určuje, že první druh obsadí část k ($k \in (0,5; 1,0)$) celkové niky, druhý druh část k' zbytku, třetí druh část k'' toho, co zbude po umístění prvního a druhého druhu, atd. do umístění všech druhů v dané nice (popis Tokeshi 1990). Teoreticky by tento proces mohl pokračovat do nekonečna.



Dominance preemption

- Tento model stejně jako předešlý představuje postupné dělení niky. První druh využije své abundance pro obsazení více než poloviny volné niky ($k \in (0,5;1,0)$) a ostatní druhy ovlivní tak, že budou obsazovat znovu více než polovinu volné niky (k' zbytku).



Log-series

- Vhodné zobrazení společenstva s log-series modelem

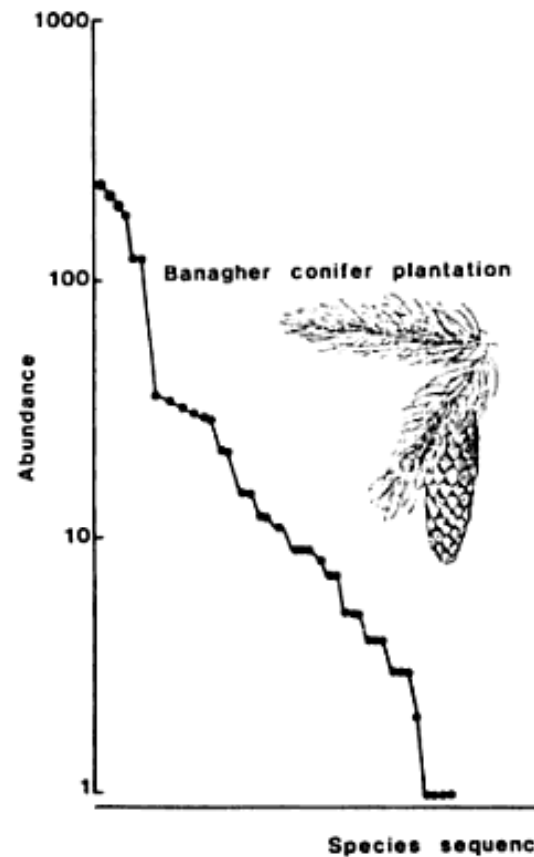
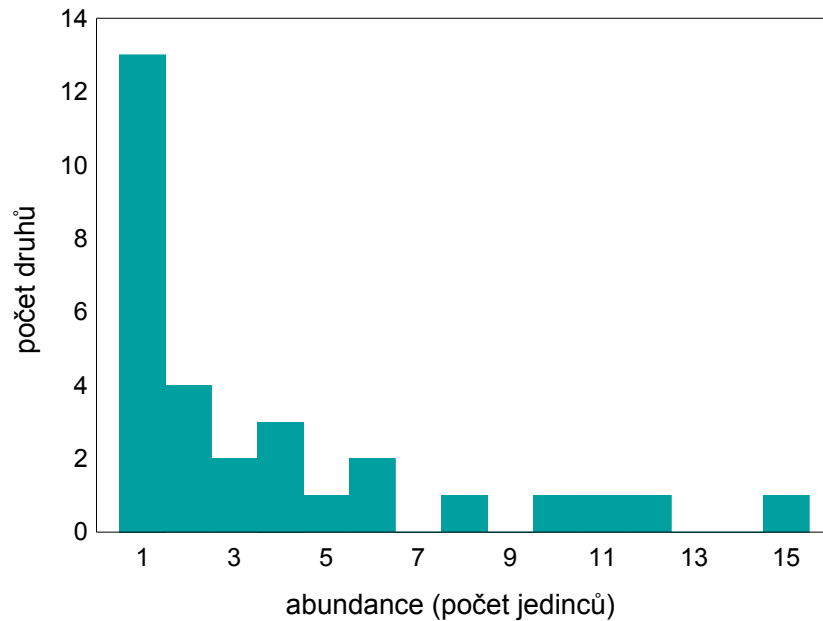


Figure 2.6 A rank abundance plot showing the diversity of ground vegetation in an Irish conifer plantation (for more information on the sites see Figure 4.2 and Chapter 4). One factor, light, has an important influence on the diversity of the vegetation, and species



Log-series

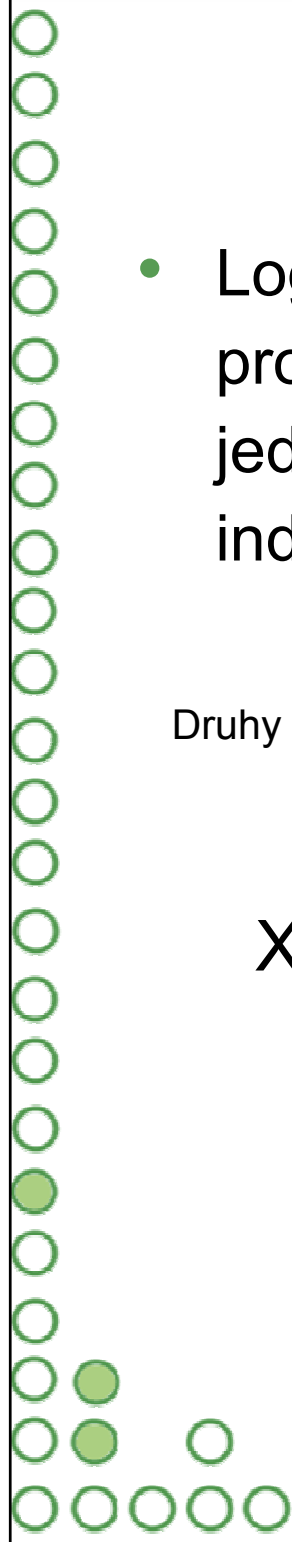
- Logaritmické řady se stejně jako geometrické řady hodí pro popis situací, kdy ekologii společenstva dominuje jeden nebo málo faktorů. Od jejich rozložení je odvozen index diverzity α .

Druhy s 1...n jedinci

$$\alpha \frac{x^1}{1}, \alpha \frac{x^2}{2}, \dots, \alpha \frac{x^n}{n} \quad S(n) = \alpha \frac{x^n}{n}$$

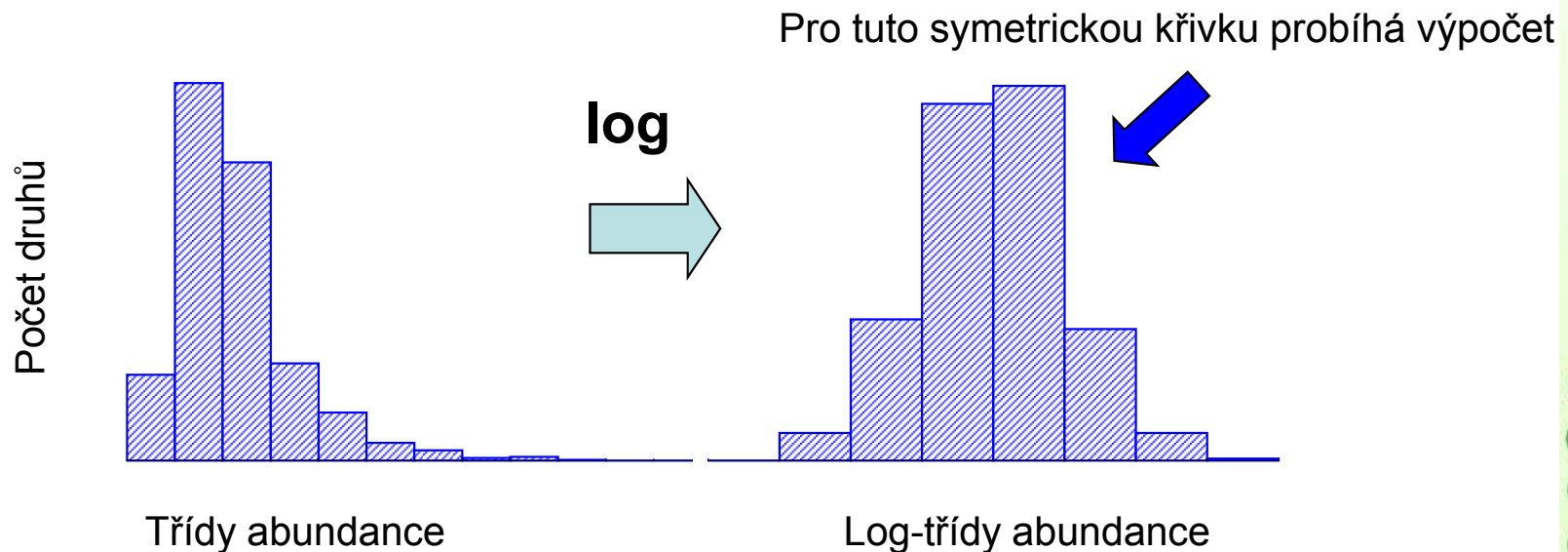
$$X : \frac{S}{N} = \frac{(1-x)[- \ln(1-x)]}{x} \quad \text{alfa : } \alpha = \frac{N(1-x)}{x}$$

$$\text{alfa variabilita: } \text{Var}(\alpha) = \frac{\alpha}{-\ln(1-x)}$$



Log normal

- Log normální rozložení je v přírodě časté – abundance, bílé krvinky
- Log normální rozložení se objevuje u mnoha společenstvech. Vyjadřuje velká, rozvinutá a pestrá přírodní společenstva. Je od něj odvozen index diverzity γ . (α bývá většinou $\sim 0,2$)



Druhy v log-třídách abundance vlevo i vpravo od středu symetrické křivky:

$$S(R) = S_0 \exp(-a^2 R^2)$$



Log normal

- Log normální rozložení by mělo být použito pouze pro spojitá data druhové abundance (pro míry jako je např. pokryvnost nebo biomasa); pro velké vzorky jsou data skoro spojitá

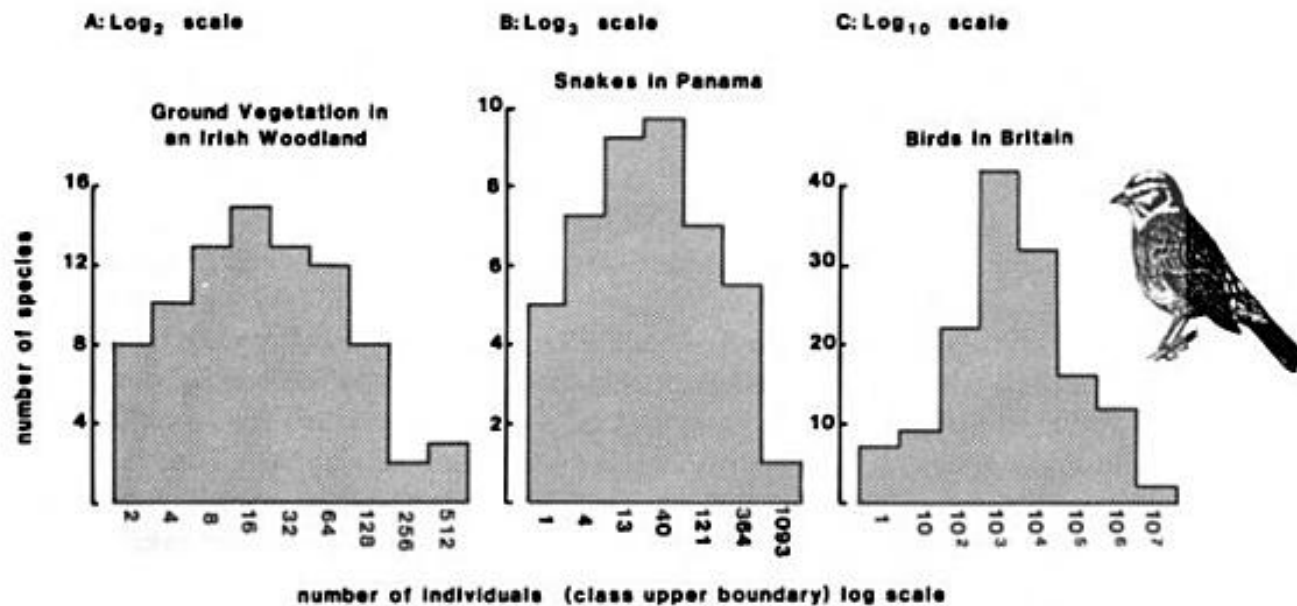
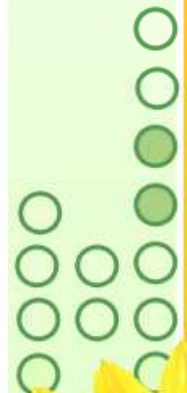


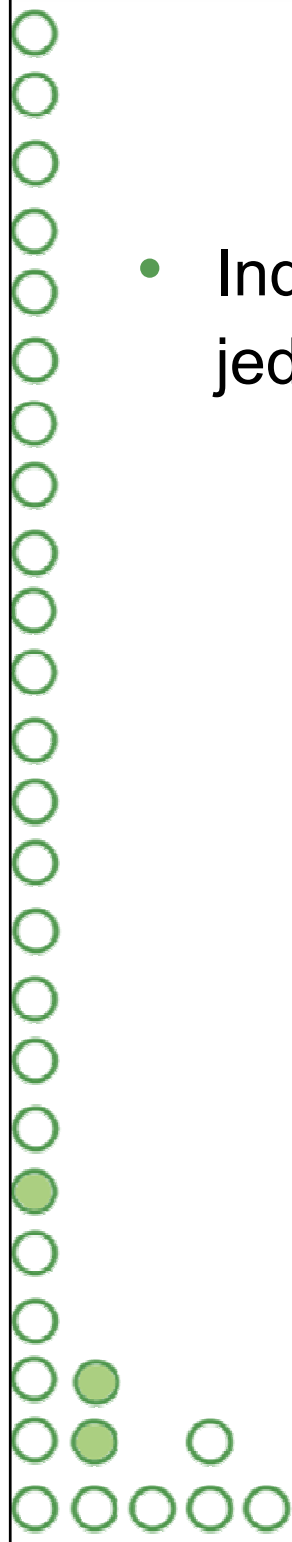
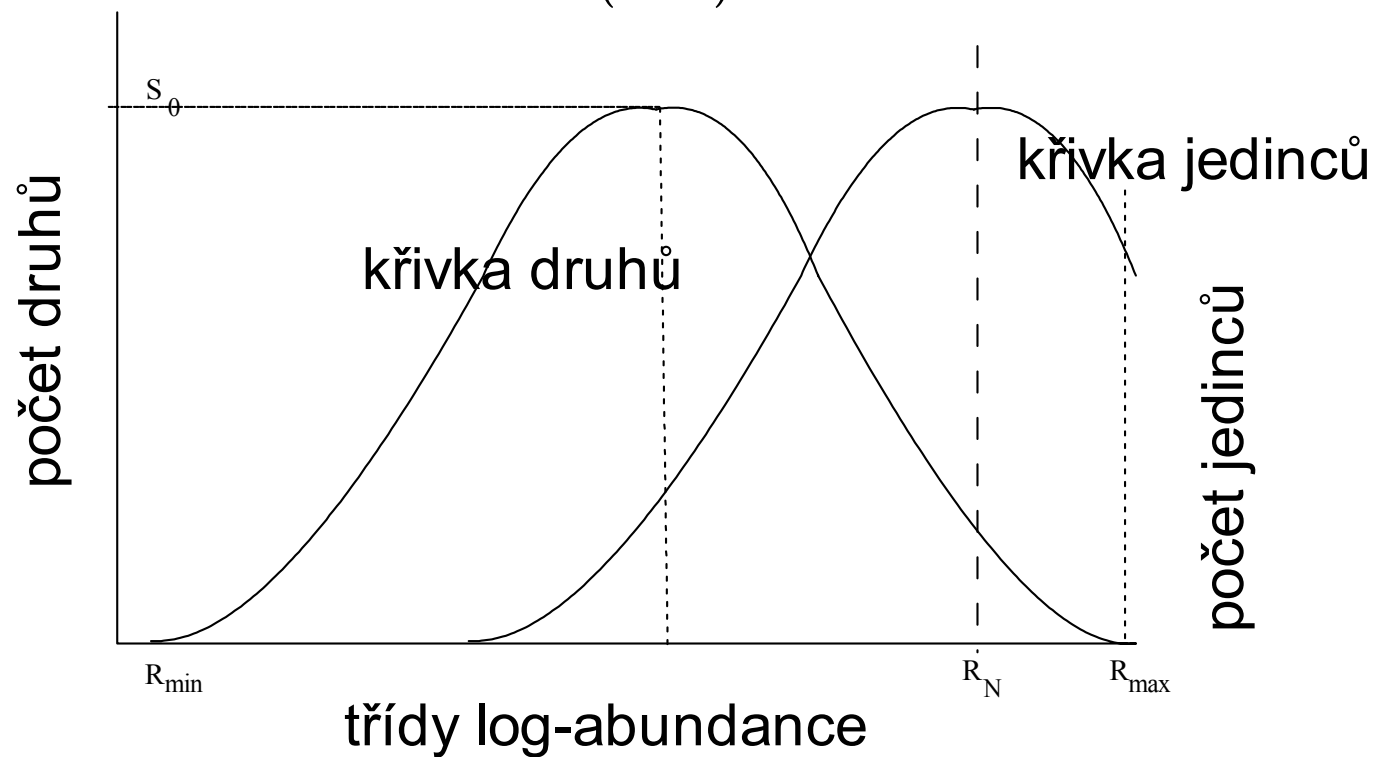
Figure 2.7 The log normal distribution I. The 'normal', symmetrical bell-shaped curve is achieved by logging the species abundances on the x-axis. A variety of log bases can be used.



γ Index diverzity

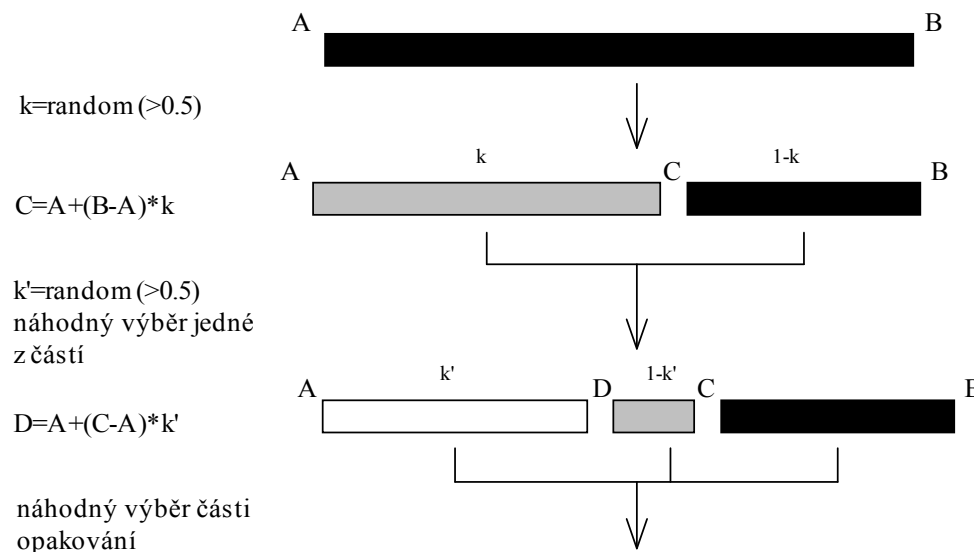
- Index diversity (γ) je mírou vztahu mezi vrcholem křivky jedinců (R_N) a vrcholem druhové křivky (R_{max}).

$$\gamma = \frac{R_N}{R_{max}} = \frac{\ln 2}{2a(\ln S_0)^{0,5}}$$



Random fraction model

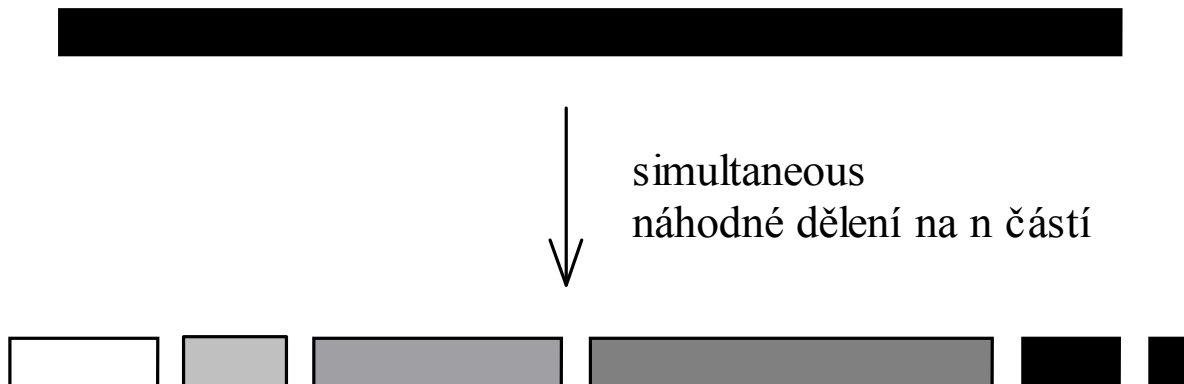
- Nejprve je celková nika náhodně rozdělena na dvě části (rovnoměrné rozložení dělicích bodů k z intervalu $(0,5;1)$). Jedna z částí je náhodně vybrána pro další dělení a opět náhodně rozdělena. Vzniklé tři části jsou opět podrobeny náhodnému výběru a z nich vybraná část náhodně rozdělena. Každá z částí má stejnou pravděpodobnost být vybrána. Tento postup představuje situaci, kdy si nově příchozí druh náhodně vybere niku jiného druhu a obsadí její náhodně velkou část. Speciálními případy jsou dominance preemption, pro dělení se vždy vybere ta nejmenší část, a dominance decay, pro dělení se vybírá největší část.



Broken stick

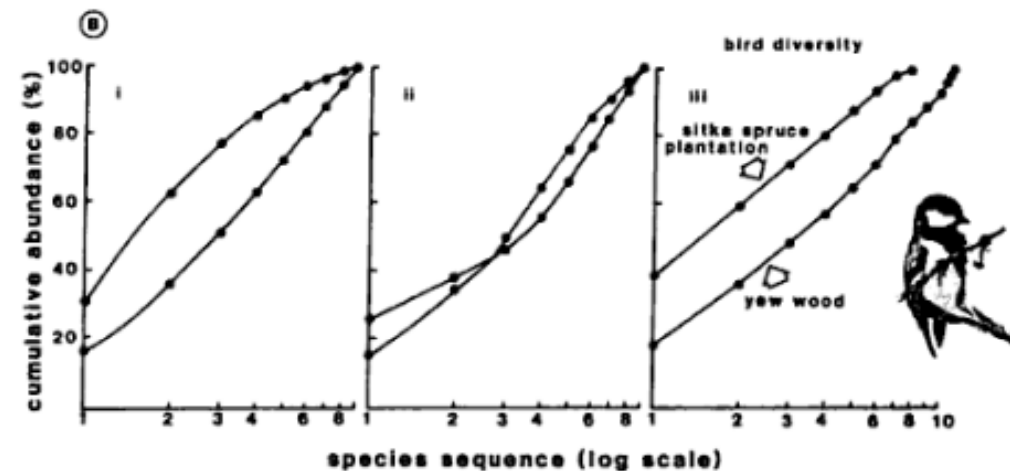
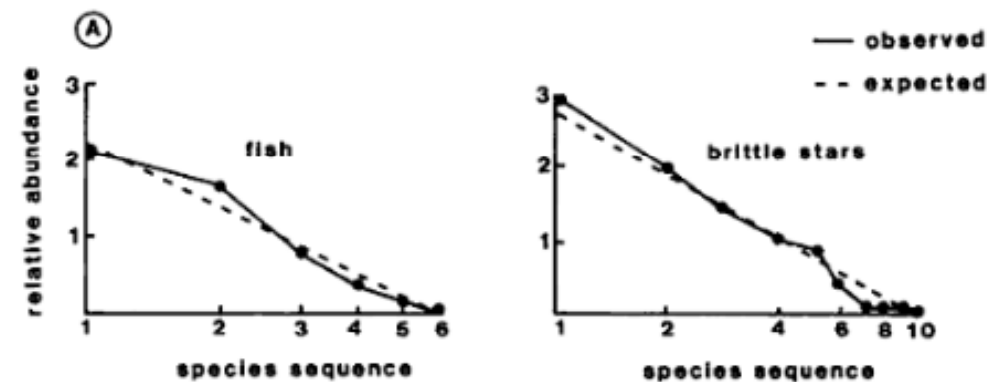
- Broken stick model odráží mnohem více vyrovnaný stav než ostatní modelová rozložení. Je to vlastně vyjádření rovnoměrného rozložení druhů. Není od něj odvozen žádný index diverzity
- Protože je charakterizován jen jediným parametrem, počtem druhů, je silně ovlivněn velikostí vzorku.

Počet druhů v kategorii s n jedinci:
$$S(n) = \frac{S(S-1)}{N} \left(1 - \frac{n}{N}\right)^{S-2}$$



Broken stick

- Zobrazení křivky společenstva s broken stick modelem: logaritmovaná pořadí druhů (A);
- K-dominance plot (B) – horní křivka patří společenstvu s nižší diverzitou



MacArthur fraction model

Tento model je shodný s broken stick modelem. Kvůli nemožnosti srovnání s ostatními modely postupného dělení niky byl původní broken stick model popsán jako model postupného dělení: Celková nika se dělí jako v případě RF modelu. Jediný rozdíl spočívá při výběru části pro další dělení. U random fraction byl tento výběr čistě náhodný, MacArthur fraction předpokládá pravděpodobnostní výběr (větší část má větší pravděpodobnost být vybrána pro dělení – je zde lineární vztah velikosti části a pravděpodobnosti jejího výběru).

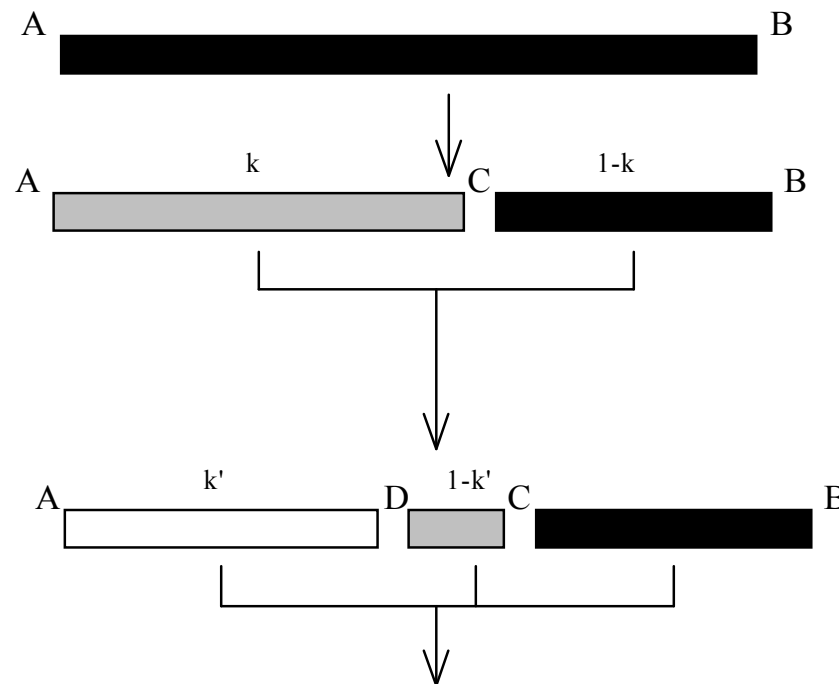
$$k = \text{random} (>0.5)$$

$$C = A + (B - A) * k$$

$k' = \text{random} (>0.5)$
výběr jedné
z částí podle
pravděpodobnosti

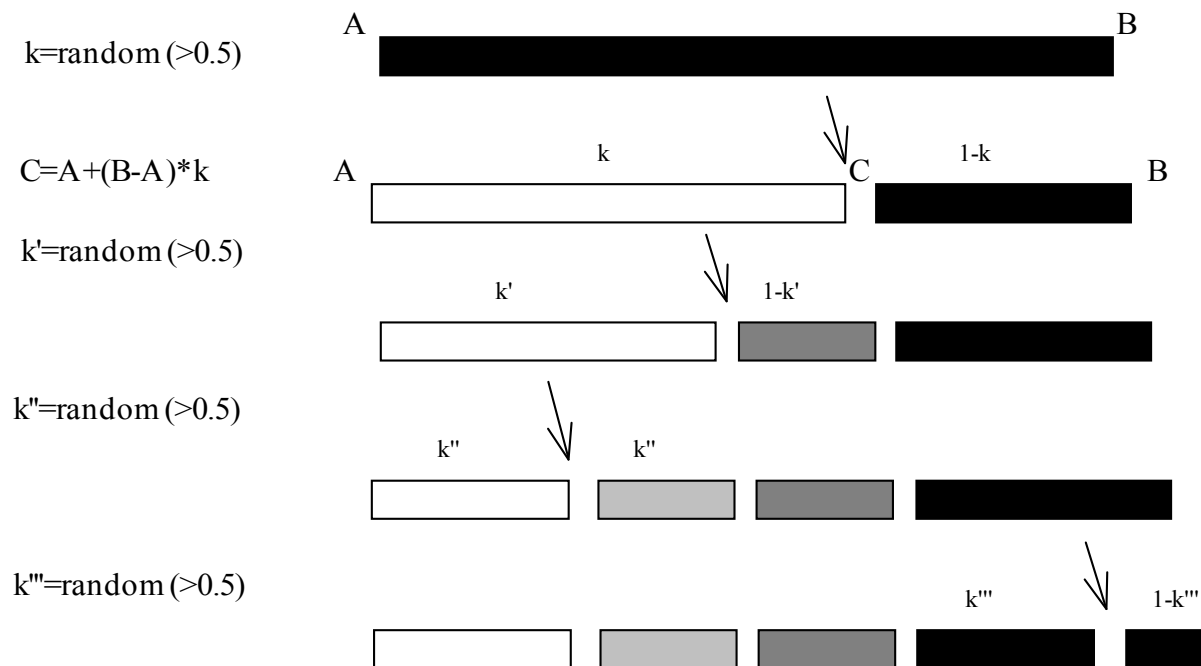
$$D = A + (C - A) * k'$$

výběr části
opakování



Dominance decay model

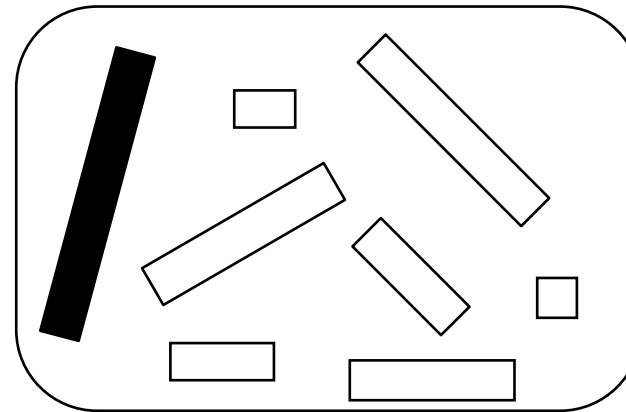
- Tento model je opakem dominance preemption modelu. Pro další dělení je vždy vybrána největší část niky, což znamená, že nově příchozí druh se vždy snaží obsadit část niky nejabundantnějšího druhu. Dominance je zde popírána. Proto výsledkem tohoto dělení jsou víceméně vyrovnané abundance jednotlivých druhů.



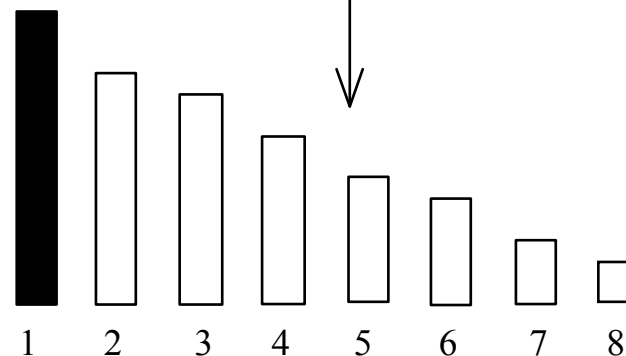
Random assortment model

- Tento model odpovídá situaci, kdy jsou abundance jednotlivých druhů na sobě nezávislé. Představuje vysoce dynamická společenstva v proměnlivém prostředí, kde nezůstává celková nika konstantní, ale její velikost se v čase mění. Pokud uvažuje soubor n nik libovolných velikostí, které uspořádáme sestupně podle velikosti, každá nika je vymezena jen svým bezprostředním větším sousedem.

nezávislé niky
různých velikostí

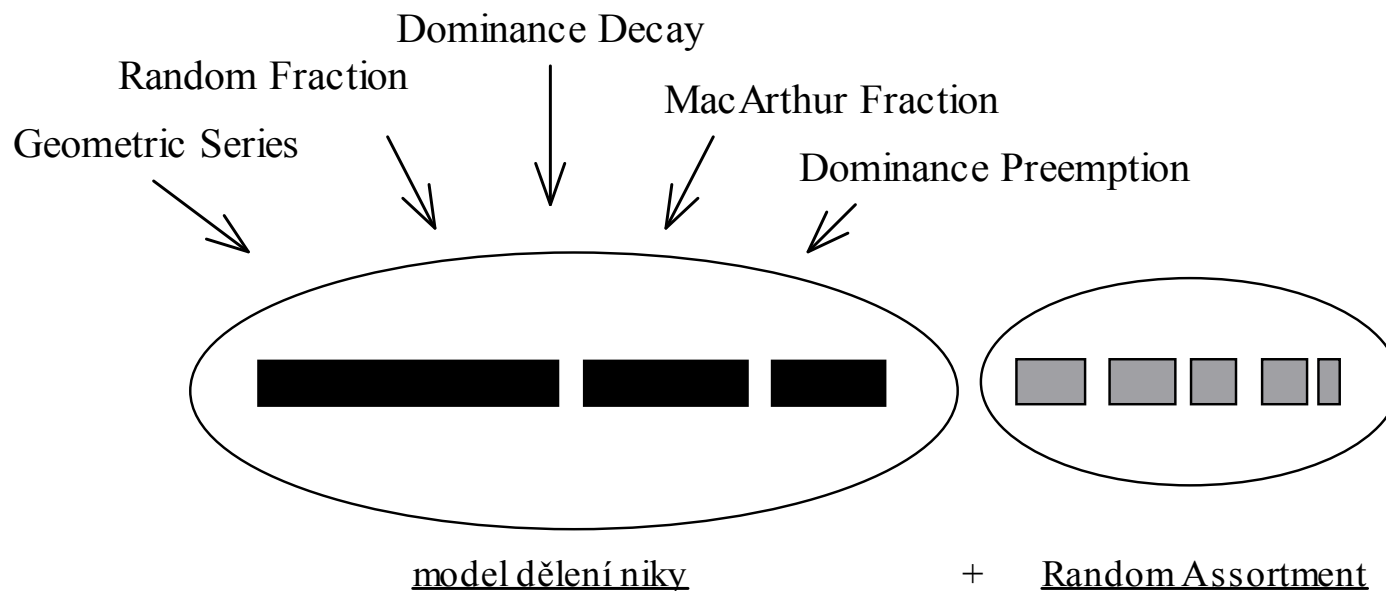


srovnání podle
velikosti



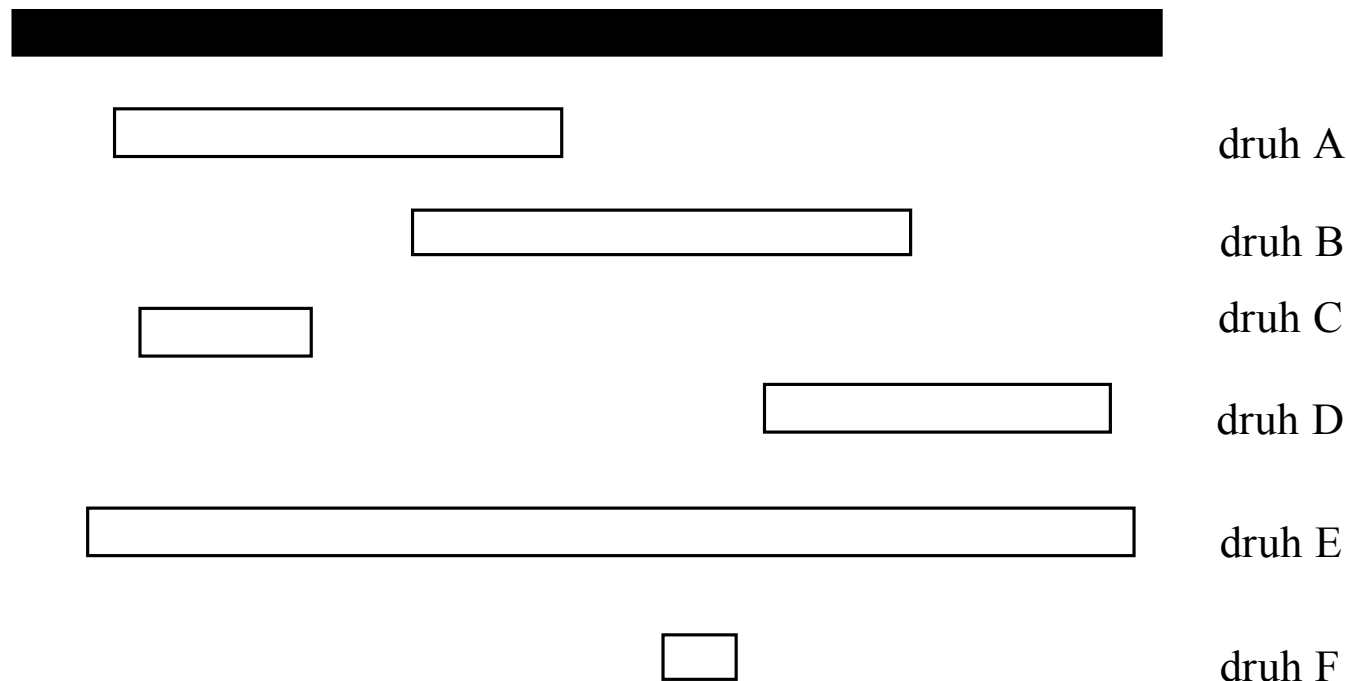
Composite model

- Tento model je odvozen od předpokladu, že společenstvo obsahuje dvě (nebo více) oddělená pravidla dělení niky. Podle Tokeshiho je pravděpodobné, že několik více abundantních druhů se řídí pravidly některého z modelů dělení niky, zatímco zbylé druhy tvoří random assortment model. Proto jednou možností, jak se dá vyjádřit, je jako spojení kteréhokoli z pěti modelů postupného dělení niky (geometric series, dominance preemption, random fraction, MacArthur fraction, dominance decay) a random assortment modelu.



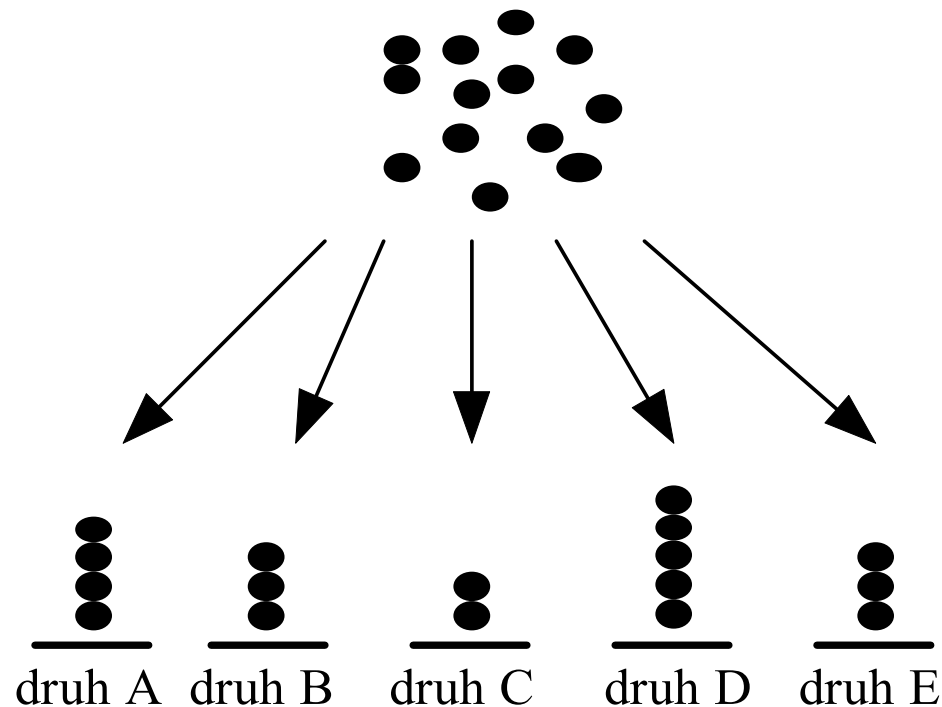
Overlapping niche model

- Niky jednotlivých druhů se mohou jakkoli překrývat, jen hranice celkové niky zůstávají stejné. Může být přiblížen k nestálému společenstvu, ve kterém vztah dvou a více druhů je výsledkem kompetice. Tento model nebyl přesně testován na reálných datech.



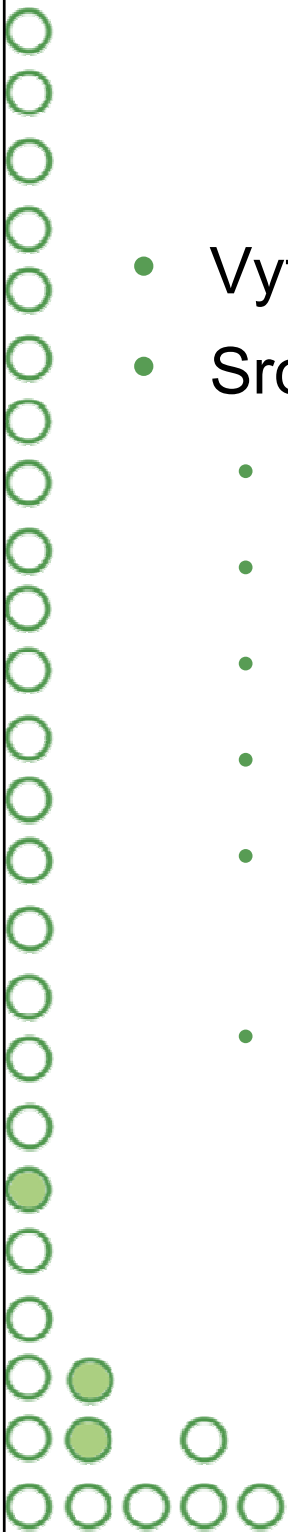
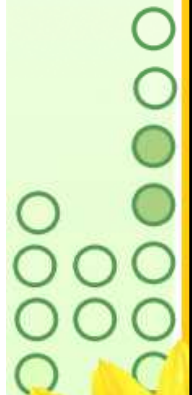
Particulate – niche model

- Podle tohoto modelu druhy ve společenstvu sdílí neměnný počet jednotek abundance („niche particles“), které jsou náhodně mezi jednotlivé druhy rozděleny. Každý z druhů má stejnou šanci na přijetí jednotky a tedy výsledkem jsou rovnoměrné abundance jednotlivých druhů. Představuje společenstvo s vrozenou rovností druhů. Tento stav teoreticky směřuje k Poissonovu rozložení druhové abundance.

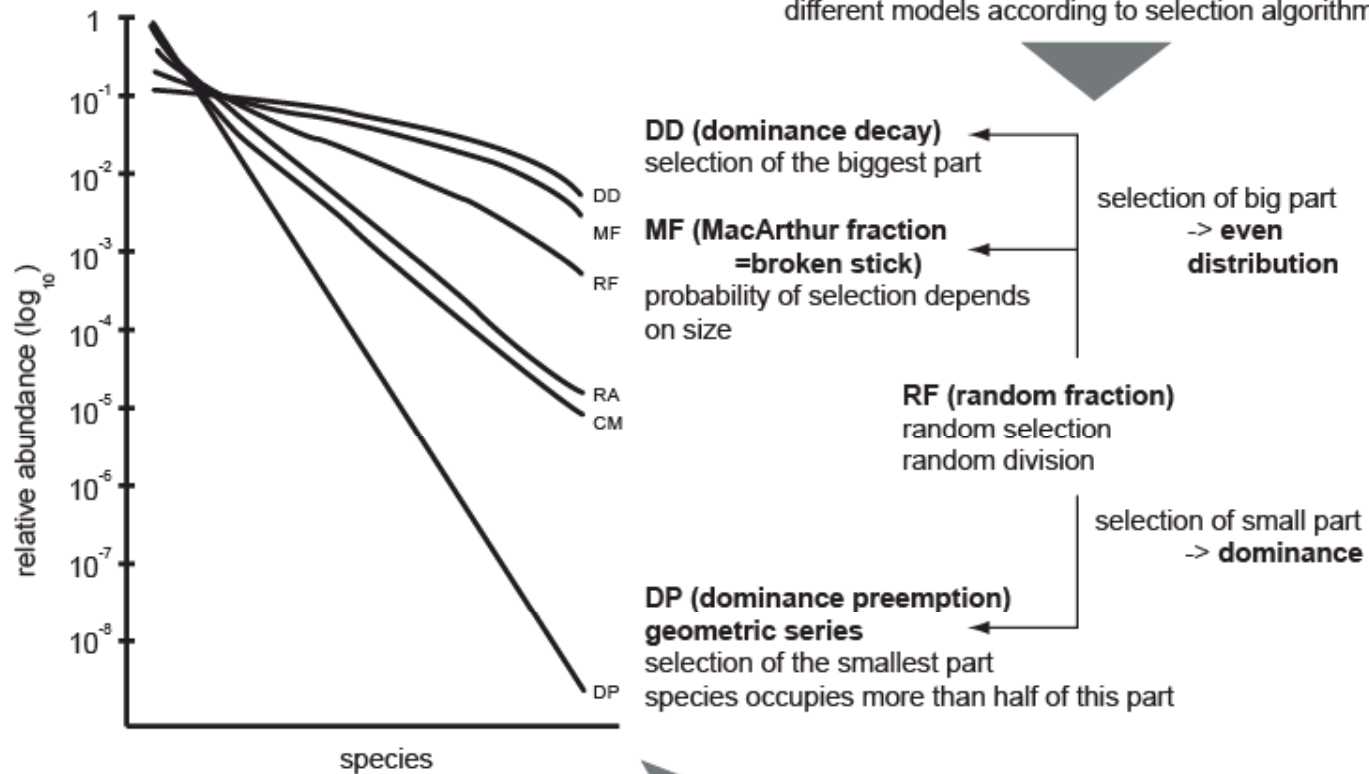
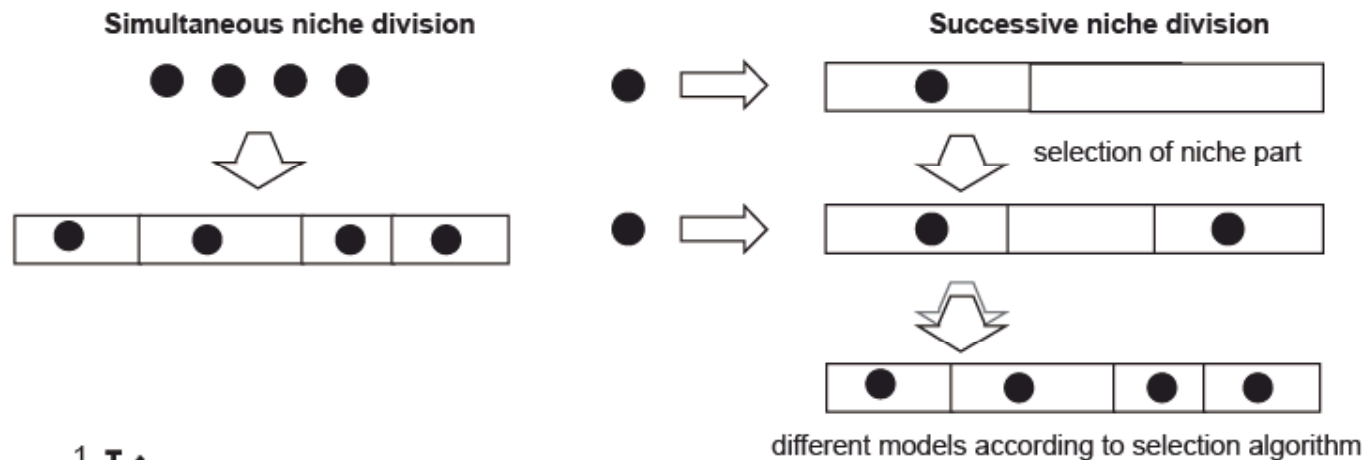


Testování modelů

- Vytvoření modelového rozložení abundancí
- Srovnání reálného a modelového rozložení
 - Opticky v grafu
 - Chi-square goodness of fit test
 - Kolgomorov - Smirnov test
 - Metriky vzdáleností
 - Srovnání s konfidenčními intervaly nasimulovaného rozložení
 - Monte Carlo testy



Na niku orientované modely: shrnutí

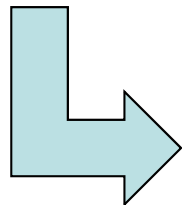


CM (composite model) = DD, MF, RF, DP for the most dominant species + RS for the rest

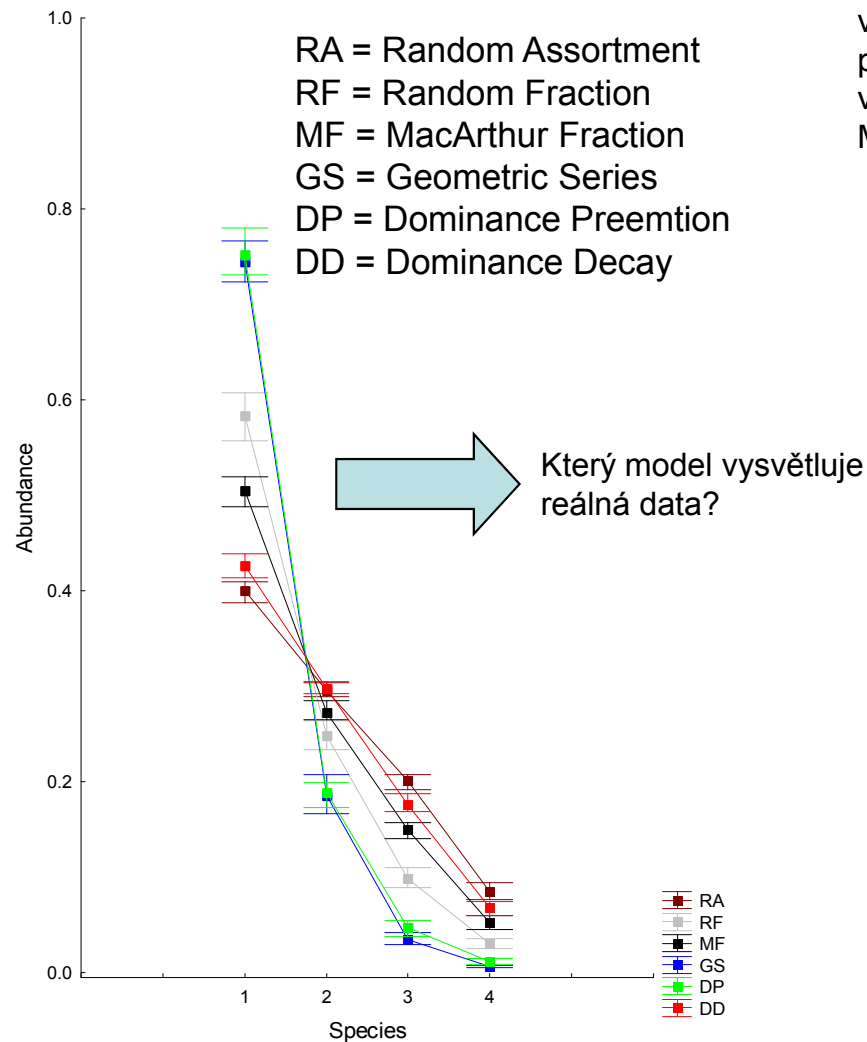
RS (random assortment) random division of niche in dynamic communities where abundance of species is independent on each other



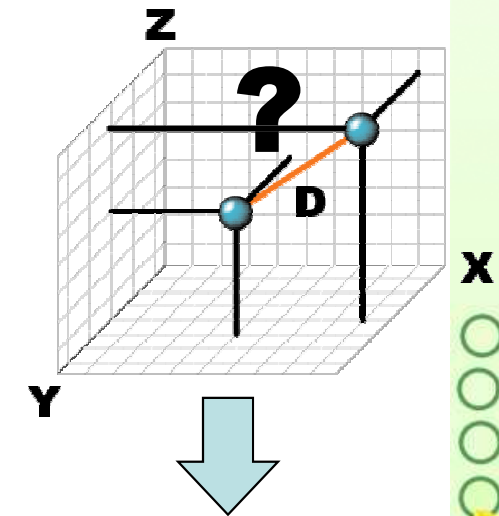
Počítačová simulace odhadu na niku orientovaných modelů navržená Tokeshi (1993)



Výpočet modelu a jeho intervalu spolehlivosti je založen na počítačové simulaci společenstva o daném počtu taxonů.



Několik metod testování vztahu mezi očekávaným a pozorovaným modelem – vícerozměrné vzdálenosti, Monte Carlo.



Závěr: data jsou vysvětlena jedním z možných na niku orientovaných modelů



Software

- Matematické modely jsou dostupné v software pro výpočet biodiverzity:
 - PAST
 - Biodiversity professional
 - Genstat
 - R project - radfit {vegan}

- Na niku orientované modely
 - Genstat
 - Software vyvinutý na IBA MU

