



# Statistické hodnocení biodiverzity

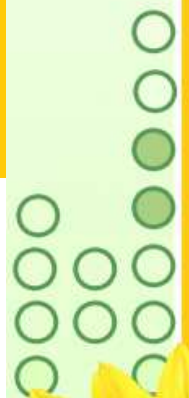
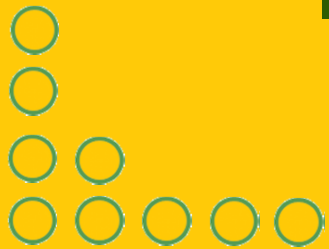
Danka Haruštiaková, Jiří Jarkovský

Institut biostatistiky a analýz, Masarykova univerzita, 2016





## IV. Indexy biodiverzity a jejich spolehlivost



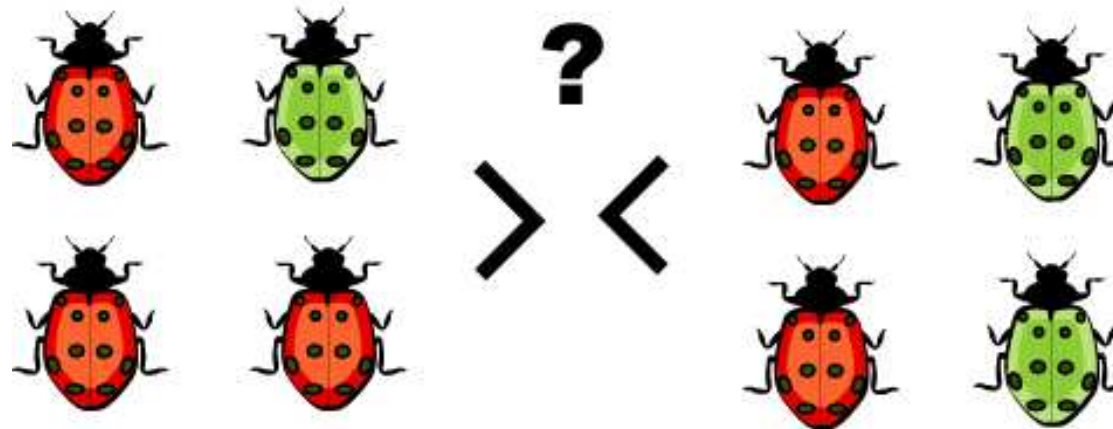
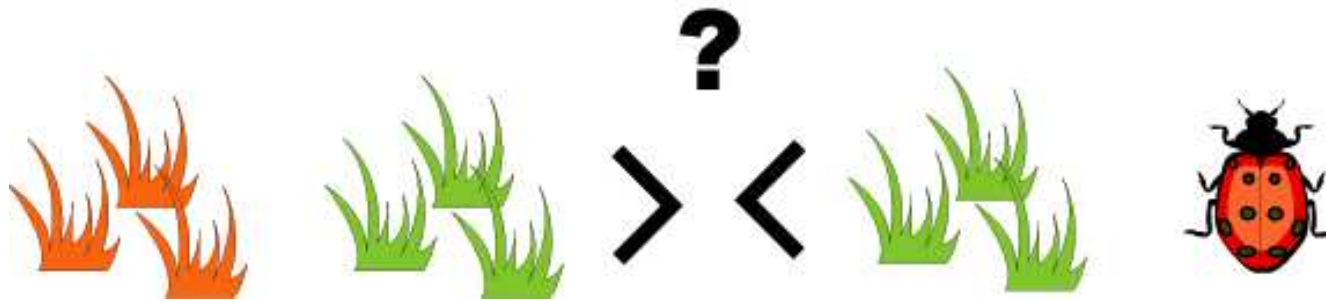
# Indexy diverzity: výhody a nevýhody

X

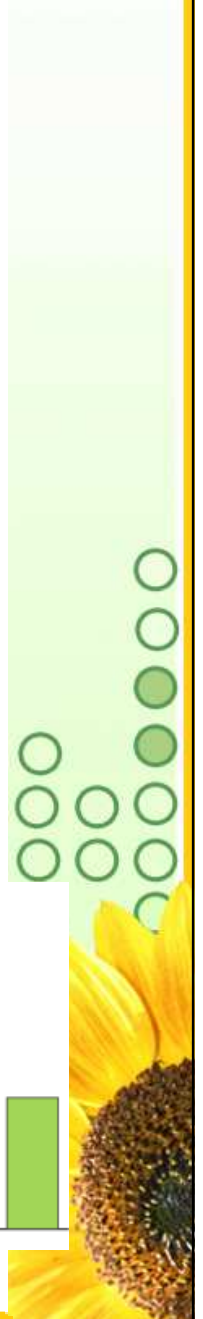
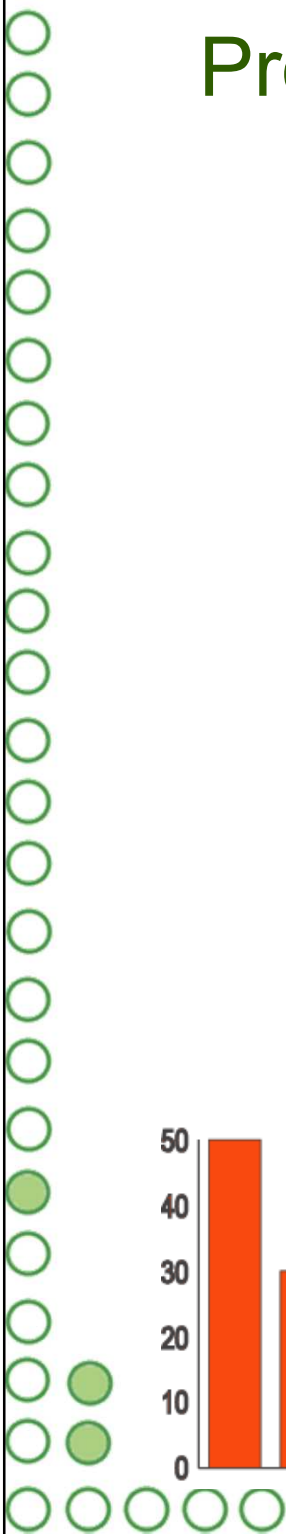
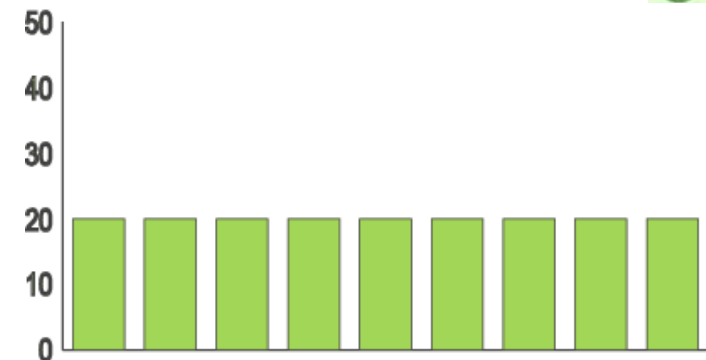
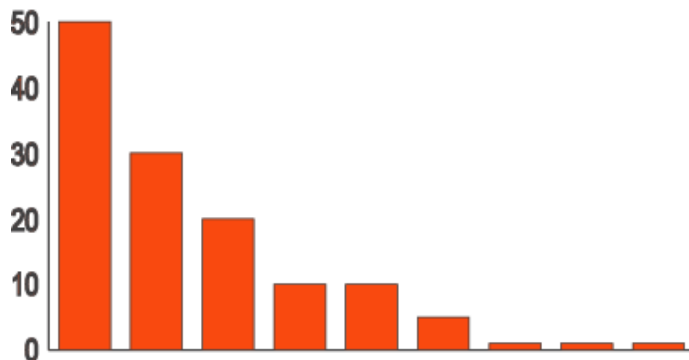
- Indexy diverzity je možné brát jako analogii k **popisné statistice**
- Celé společenstvo je agregováno **jediným číslem**, které reprezentuje **počet druhů a/nebo jejich dominanci ve společenstvu**
- Pro popisnou statistiku diverzity je možné získat intervaly spolehlivosti a dostupné jsou i statistické testy
- **Výhody:**
  - Měření diverzity v jediném čísle
- **Nevýhody:**
  - Redukce individuality taxonů
  - V některých případech nejasná interpretace (stejná hodnota indexu může být spočítána z velmi odlišných společenstev)



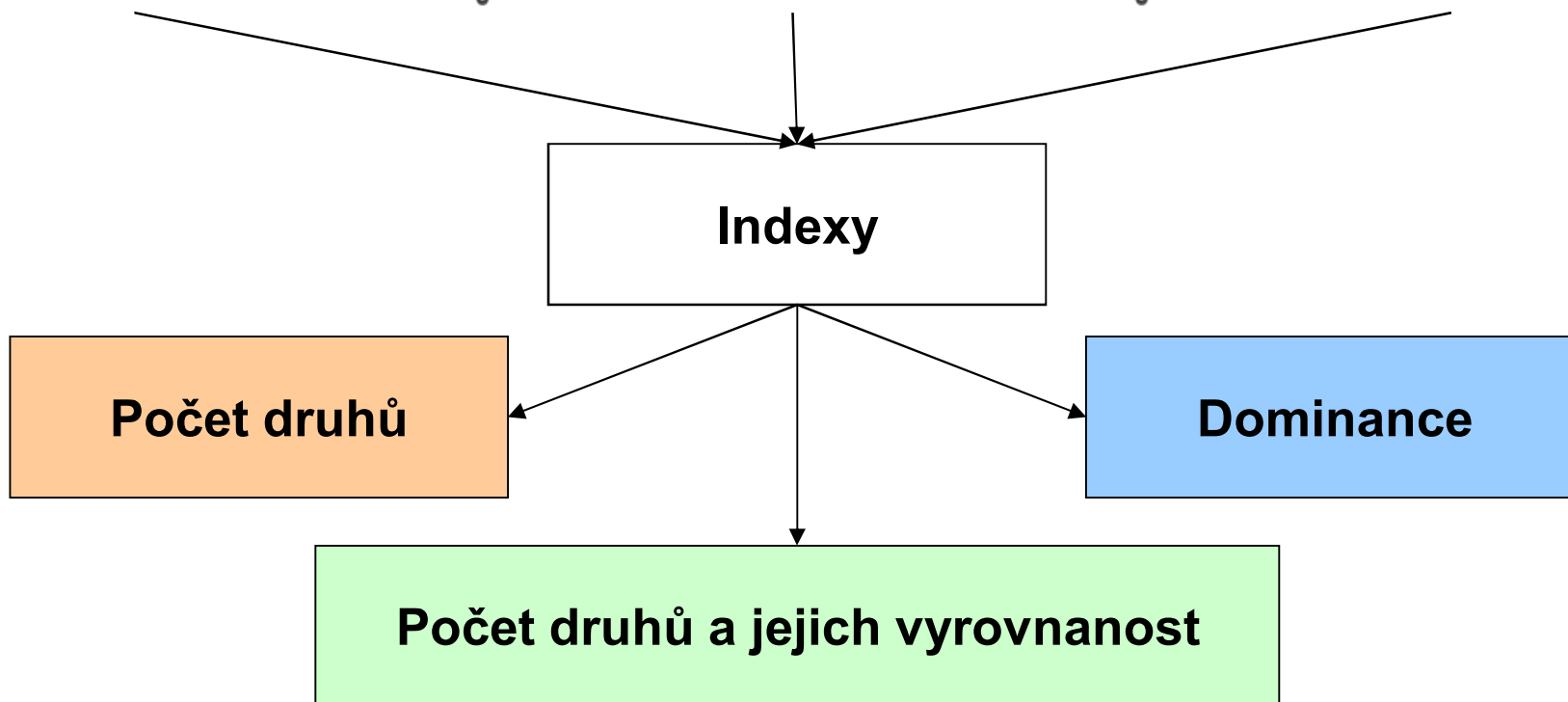
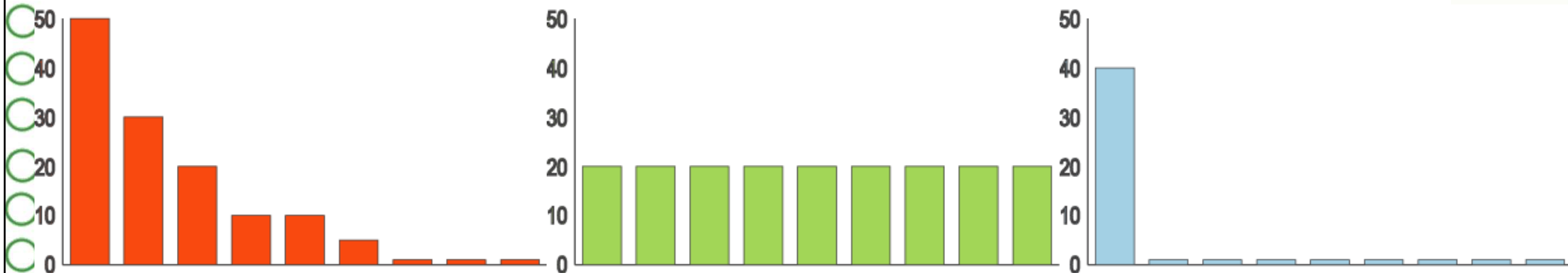
# Problémy s biodiverzitou – co znamená větší diverzitu ?



?



# Indexové hodnocení diverzity





# Indexy diverzity

$S$  'species richness'

$$R1 = (S - 1) / \ln N$$

$$R2 = S / \sqrt{N}$$

$$H' = - \sum p_i \ln p_i$$

$H'_{adj}$

$$N1 = e^{H'}$$

$$Q = \frac{\frac{1}{2} n_{R1} + \sum_{R1+1}^{R2-1} n_r + \frac{1}{2} n_{R2}}{\log\left(\frac{R2}{R1}\right)}$$

$$HB = \frac{\ln N! - \sum \ln n_i!}{N}$$

$$d = \frac{N_{max}}{N}$$

*McIntosh, 1967*

*Margalef, 1958*

*Menhinick, 1964*

*Shannon & Weaver, 1949*

*Hutcheson, 1970*

*Simpson, 1949*

*Kempton & Taylor, 1978*

*Brillouin, 1956*

*Berger & Parker, 1970*



# Indexy druhové bohatosti

- Počet druhů (**druhová bohatost**)  $S$
- Počet druhů vztažený na počet jedinců nebo biomasu  
v případě, že není možné zajistit stejnou velikost vzorku (počet druhů narůstá s velikostí vzorku a vzorkovacím úsilím) => RAREFACTION – metoda výpočtu počtu druhů očekávaných ve vzorku standardní velikosti (např. 1000 jedinců)
- Hustota druhů – počet druhů v dané oblasti

Margalefův index  $R1 = (S - 1) / \ln N$

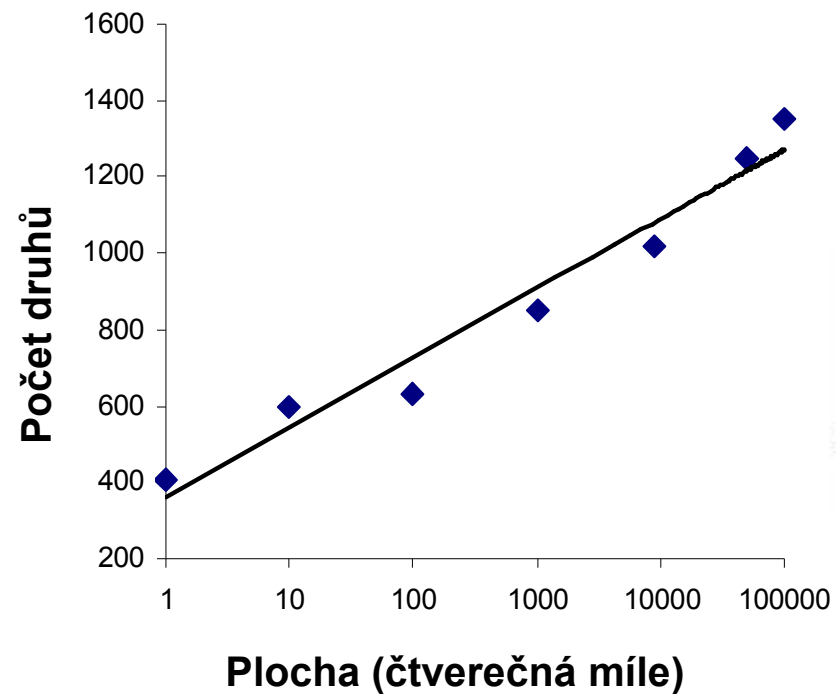
Menhinickův index  $R2 = S / \sqrt{N}$

Počet druhů je vážen počtem jedinců.

- **Nevýhody indexů druhové bohatosti:** křivka společenstva není uvažována ve výpočtu

# Druhová bohatost

- Druhová bohatost narůstá s velikostí vzorku.



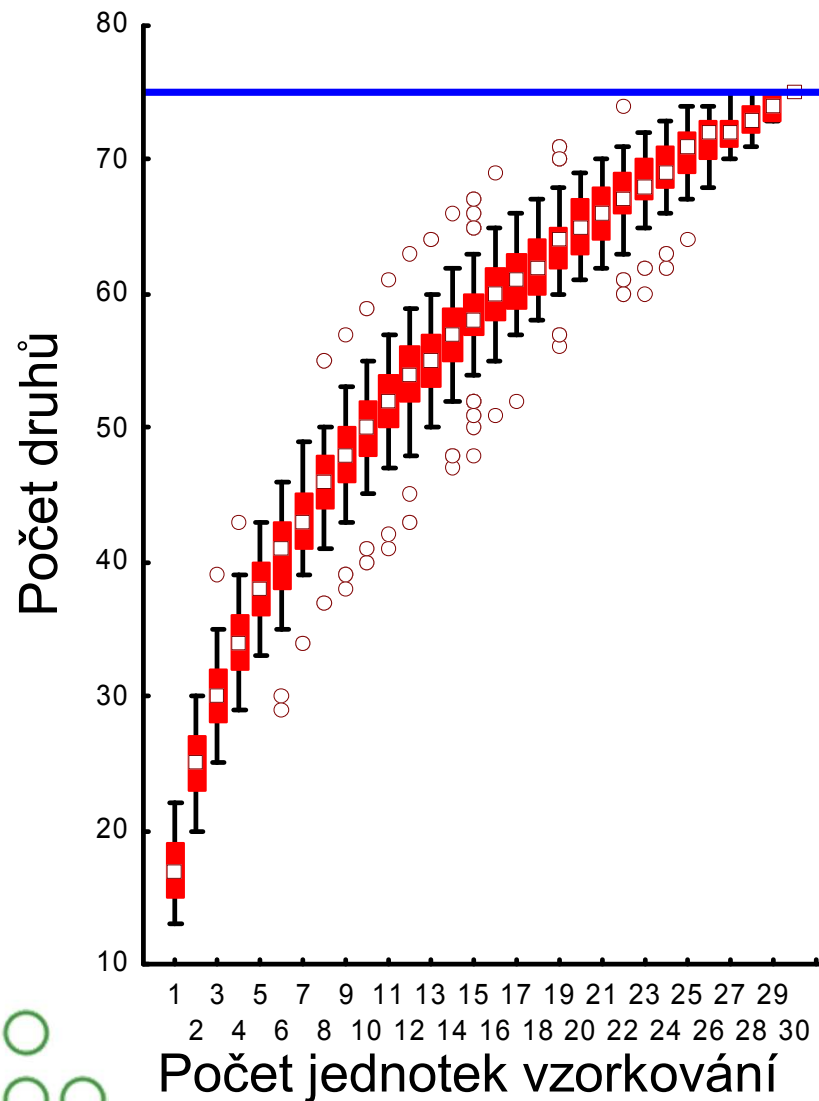
Relationship between number of species and area for flowering plants in England.  
Redrawn from Krebs (1985) after Williams (1964).





# Velikost vzorku a počet druhů

- Počet druhů nelineárně závisí na počtu jedinců ve vzorku



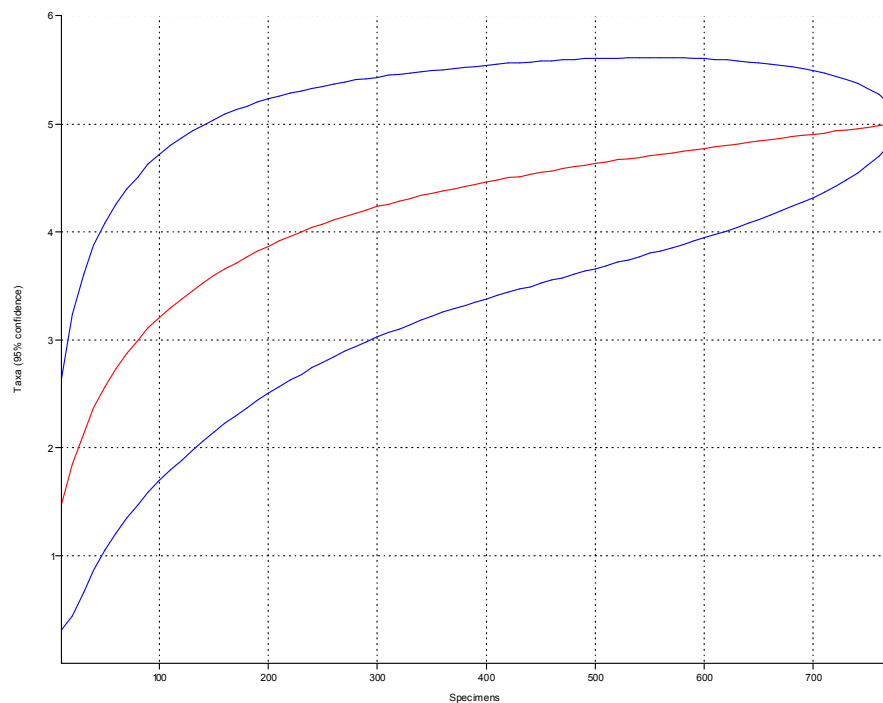
- Jak srovnat druhovou bohatost ve vzorku různé velikosti?

RAREFACTION



# Rarefaction

- Metoda řešící problém srovnání druhové bohatosti ve vzorcích o různé velikosti
- Standardizuje oba vzorky na velikost menšího vzorku sample



Metoda výpočtu odhadu počtu **druhů v náhodném výběru o n jedincích ze vzorku**

Je možno spočítat jak bodové, tak intervalové odhady



# Rarefaction: výpočet

$$E(\hat{S}_n) = \sum_{i=1}^s \left[ 1 - \frac{\binom{N - N_i}{n}}{\binom{N}{n}} \right]$$

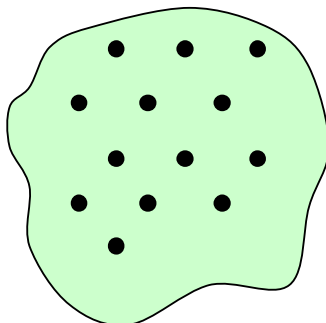
- Kde **E(S<sub>n</sub>)** je odhadnutý počet druhů ve vzorku o **n** jedincích, **N** je celkový počet jedinců, **N<sub>i</sub>** je počet jedinců druhu **i** a **n** je počet jedinců pro které je odhad počítán.
- Variabilita odhadu je vypočtena:

$$\text{var}(\hat{S}_n) = \binom{N}{n}^{-1} \left[ \sum_{i=1}^s \binom{N - N_i}{n} \left[ 1 - \frac{\binom{N - N_i}{n}}{\binom{N}{n}} \right] + 2 \sum_{i=1}^{s-1} \sum_{j=i+1}^s \left[ \binom{N - N_i - N_j}{n} - \frac{\binom{N - N_i}{n} \binom{N - N_j}{n}}{\binom{N}{n}} \right] \right]$$

# Rarefaction: předpoklady

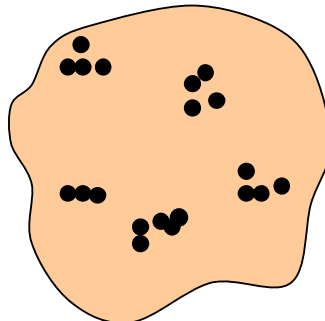
- Společenstva porovnávaná pomocí rarefaction mají mít podobné taxonomické složení
- Srovnatelné metody vzorkování
- Rarefaction křivka nemůže být extrapolována za sumu jedinců největšího vzorku
- Jedinci jsou ve společenstvu náhodně rozmístěni
  - Ve skutečnosti jsou častější shluky jedinců a za těchto podmínek dává rarefaction nadsazené odhady

$$\sigma^2 < \mu$$



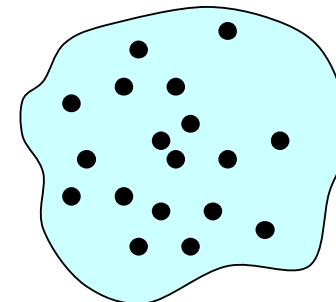
Uniform

$$\sigma^2 > \mu$$

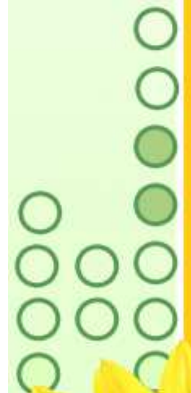


Clustered

$$\sigma^2 = \mu$$

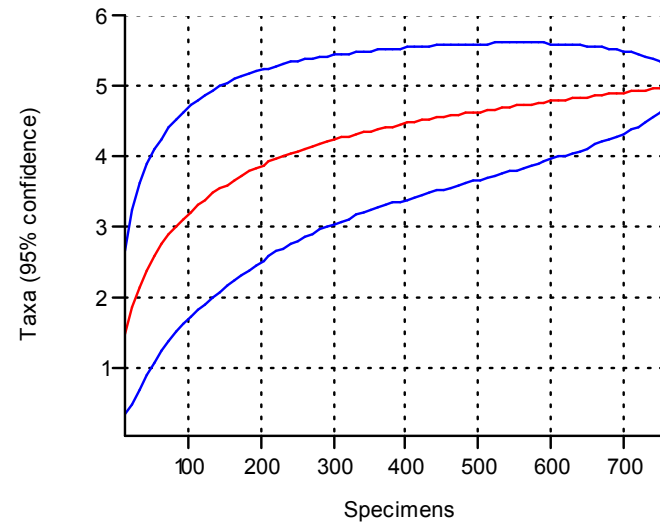


Random

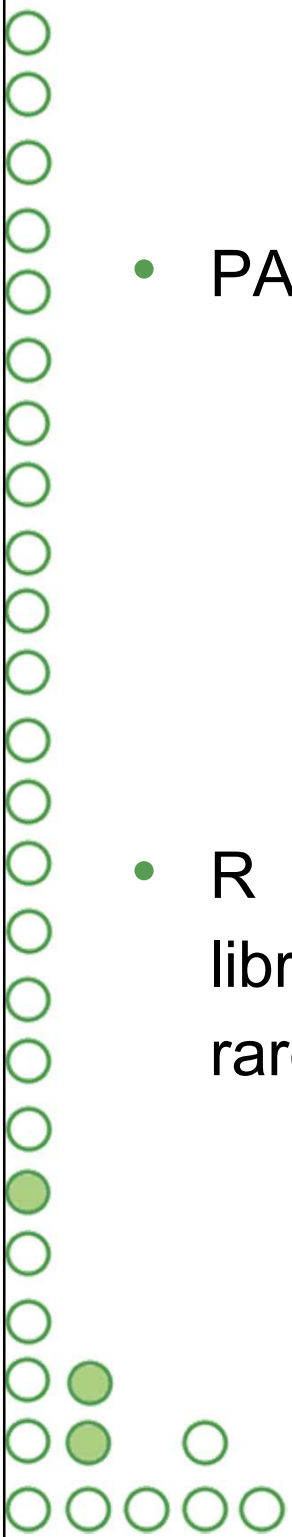


# Software

- PAST software



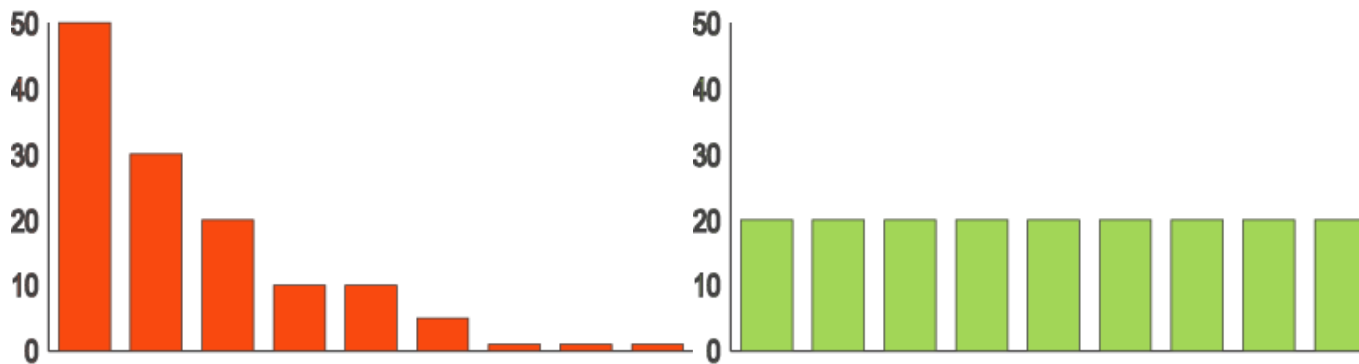
- R  
library(vegan)  
rarefy





# Vyrovnanost společenstva a počet druhů

- Doplnkem těchto indexů je tzv. evenness, která je počítána jako podíl indexu skutečného společenstva a teoretické maximální hodnoty pro daný počet druhů (tj. pokud by měly zcela vyrovnané abundance) – evenness vlastně vyjadřuje jak moc je reálné společenstvo vzdáleno od maximální vyrovnanosti



- Do této skupiny patří Shannonův a Brillouinův index, liší se od sebe použitím, Brillouinův index by měl být používán pouze pro skutečně vzorkovanou část společenstva (příkladem mohou být infrakomunity parazitů, kdy sesbíráme všechny parazity na hostiteli), Shannonův index při výpočtu uvažuje, že část společenstva vzorkována nebyla



# Shannon index

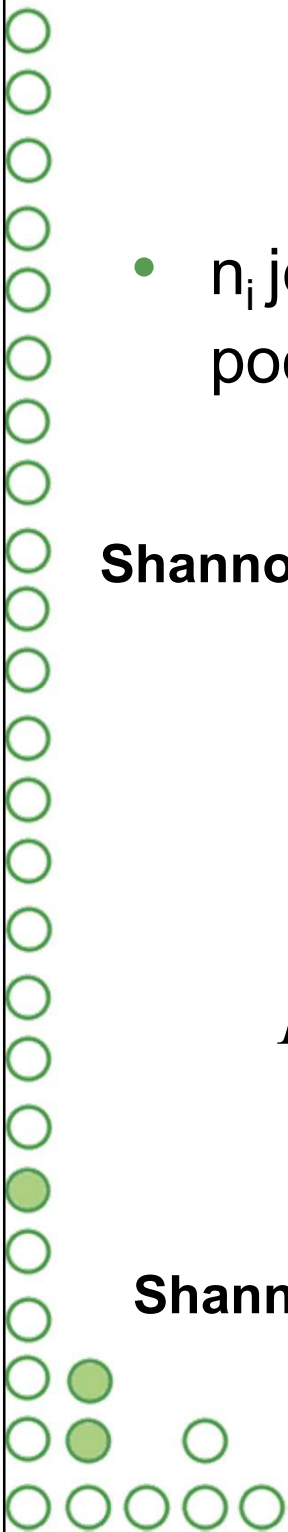
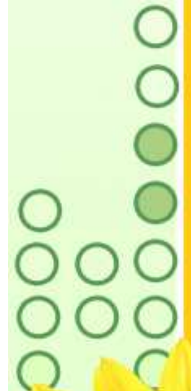
- $n_i$  je abundance i-tého druhu,  $N$  celkový počet jedinců a  $S$  počet druhů

**Shannonův index**  $H' = -\sum p_i \ln p_i$        $p_i = \frac{n_i}{N}$

## Shannonův index s korekcemi

$$H' = -\sum p_i \ln p_i - \frac{S-1}{N} + \frac{1 - \sum p_i^{-1}}{12N^2} + \frac{\sum (p_i^{-1} - p_i^{-2})}{12N^3}$$

**Shannon evenness**  $E = \frac{H'}{H_{\max}} = \frac{H'}{\ln S}$





# Shannon index

## Shannon index

$$H' = -\sum p_i \ln p_i \quad p_i = \frac{n_i}{N}$$

$n_i$  ... abundance  $i$ -tého druhu

$N$  ... celkový počet jedinců

Species	Abundance	Shannon index of diversity		
		$p_i$	$\ln(p_i)$	$p_i * \ln(p_i)$
 Baetis alpinus	736	0.9472	-0.0542	-0.0513
Rhithrogena semicolorata	28	0.0360	-3.3232	-0.1198
Epeorus sylvicola	8	0.0103	-4.5760	-0.0471
 Baetis rhodani	4	0.0051	-5.2691	-0.0271
Ephemerella mucronata	1	0.0013	-6.6554	-0.0086
<b>Total number of individuals</b>	<b>777</b>			
<b>Shannon index of diversity (using natural logarithm)</b>				<b>0.2539</b>

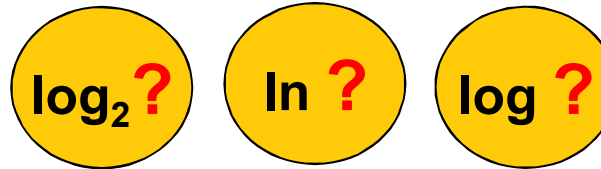
Shannon index je mezi ekology nejpopulárnější index.





# Shannon index

Shannon index

$$H' = -\sum p_i \ln p_i \quad p_i = \frac{n_i}{N}$$



Species	Abundance	Shannon index of diversity		
		$p_i * \log_2(p_i)$	$p_i * \ln(p_i)$	$p_i * \log(p_i)$
 Baetis alpinus	736	-0.0741	-0.0513	-0.0223
Rhithrogena semicolorata	28	-0.1728	-0.1198	-0.0520
Epeorus sylvicola	8	-0.0680	-0.0471	-0.0205
 Baetis rhodani	4	-0.0391	-0.0271	-0.0118
Ephemerella mucronata	1	-0.0124	-0.0086	-0.0037
<b>Total number of individuals</b>	<b>777</b>			
<b>Shannon index of diversity</b>		<b>0.3663</b>	<b>0.2539</b>	<b>0.1103</b>

Báze logaritmu



Shannon index



# Shannon index

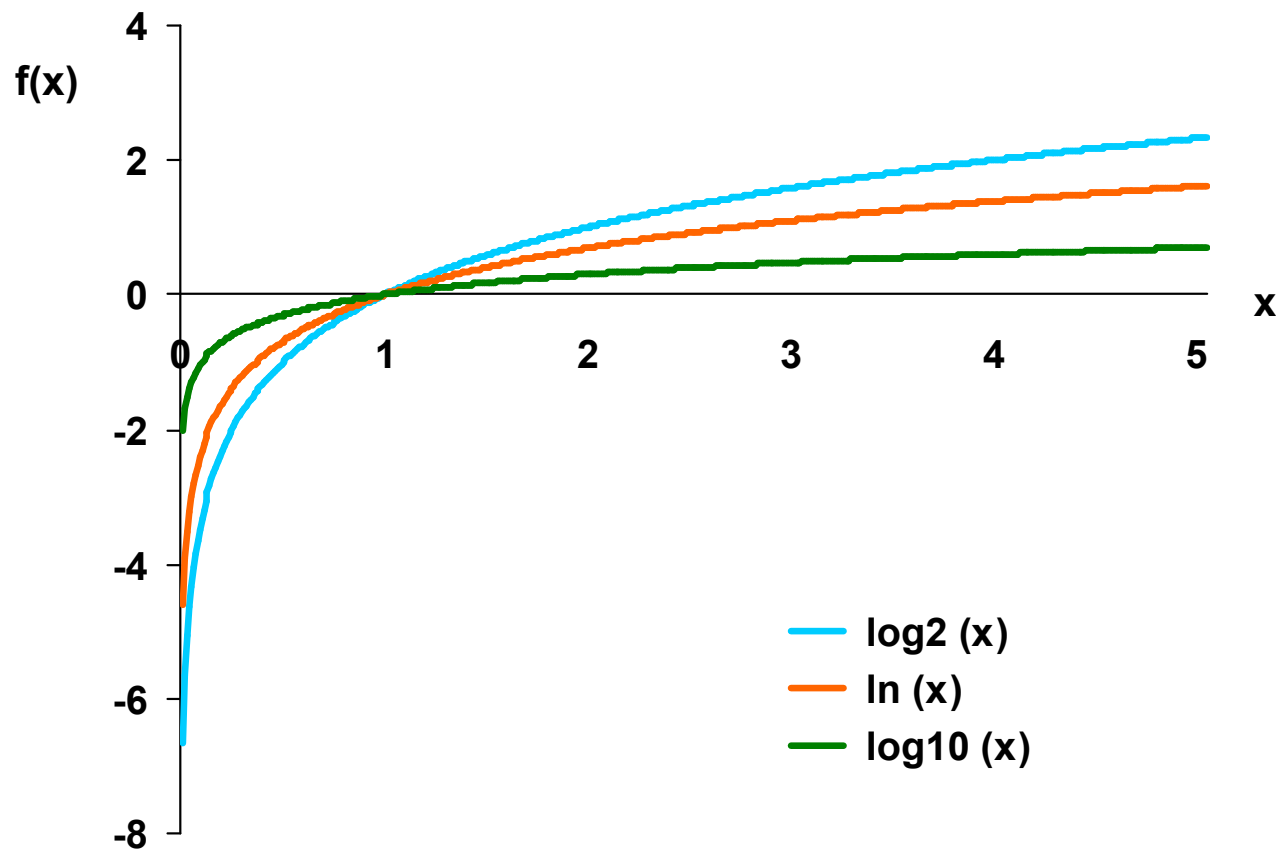
Shannon index

$$H' = -\sum p_i \ln p_i \quad p_i = \frac{n_i}{N}$$

Báze logaritmu



Shannon index





# Brillouinův index

- Možným problémem při výpočtu je neschopnost Excelu spočítat faktoriál více než 160, lze obejít použitím logaritmu

**Brillouinův index**

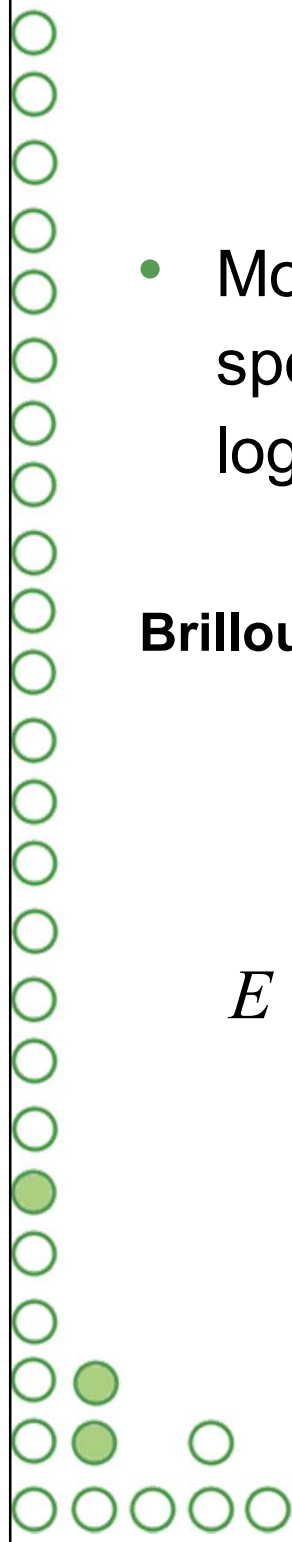
$$HB = \frac{\ln N! - \sum \ln n_i!}{N}$$

## Brillouin evenness

$$E = \frac{HB'}{HB_{\max}}$$

$$HB_{\max} = \frac{1}{N} \ln \frac{N!}{\left( \left[ \frac{N}{S} \right]! \right)^{s-r} \left( \left( \left[ \frac{N}{S} \right] + 1 \right)! \right)^r}$$

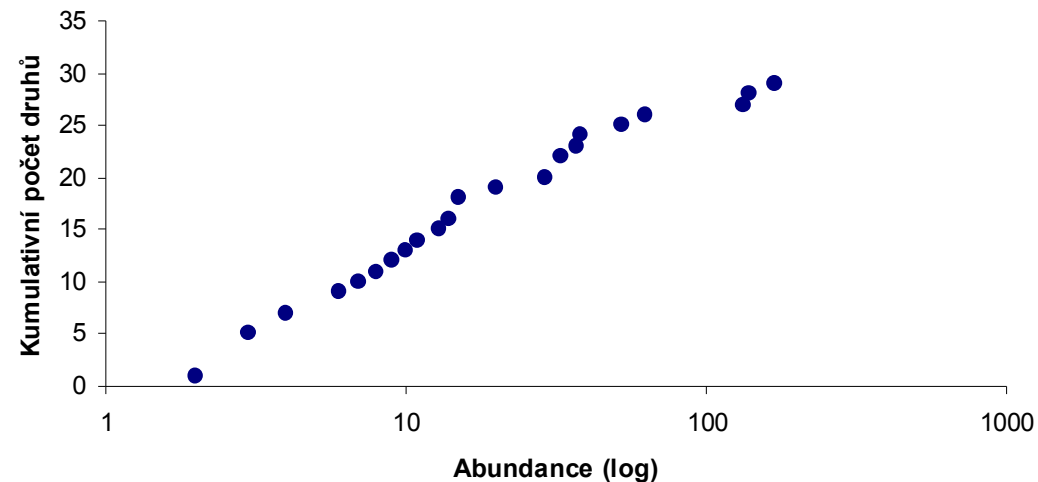
kde  $\left[ \frac{N}{S} \right]$  je celá část  $\frac{N}{S}$  a  $r = N - S \left[ \frac{N}{S} \right]$



# Q statistic I

- Měřítka sklonu křivky abundancí kumulativního počtu druhů

$$Q = \frac{\frac{1}{2}n_{R1} + \sum_{R1+1}^{R2-1} n_r + \frac{1}{2}n_{R2}}{\log\left(\frac{R2}{R1}\right)}$$



$\sum n_r$  – celkový počet druhů mezi kvartily

S – celkový počet druhů ve vzorku

R1 a R2 – 25% a 75% kvartil

$n_{R1}$  – počet druhů ve třídě, do níž spadá dolní kvartil počtu druhů

$n_{R2}$  – počet druhů ve třídě, do níž spadá horní kvartil počtu druhů

R1 – počet jedinců ve třídě, do níž spadá dolní kvartil počtu druhů

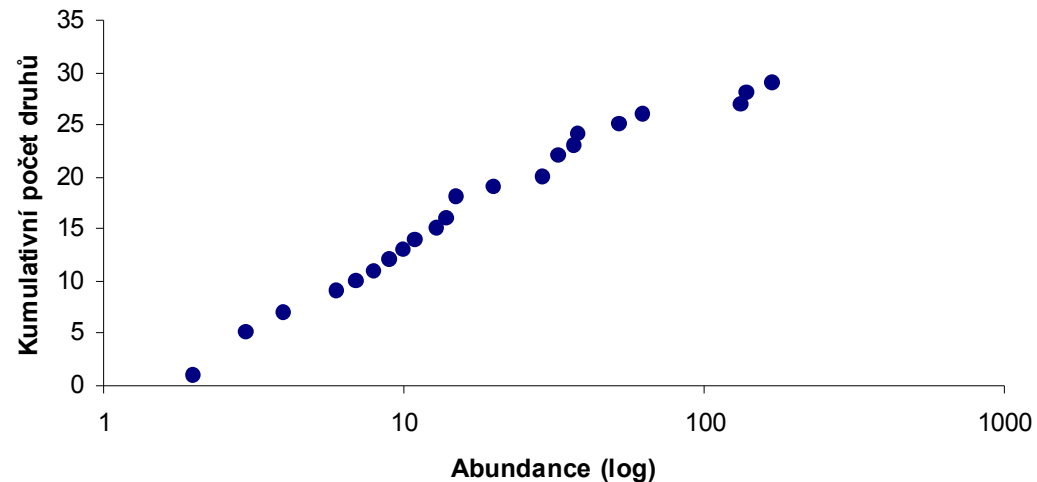
R2 – počet jedinců ve třídě, do níž spadá horní kvartil počtu druhů



# Q statistic II

- Dalším způsobem výpočtu je odhad hodnoty Q

$$X_i = \frac{S_j - S_{j'}}{\log\left(\frac{N_j}{N_{j'}}\right)}$$



je počítán pro všechny páry  $S_j$  a  $S_{j'}$  a  $N_j$  a  $N_{j'}$  ( $j > j'$ ,  $j=1,2, \dots, r$ )

S – kumulativní počet druhů

N – počet jedinců ve třídě

r – počet tříd a  $(i=1,2, \dots, r(r-1)/2)$

Z vzniklé řady čísel je Q zjištěno jako medián nebo geometrický průměr



# Dominance

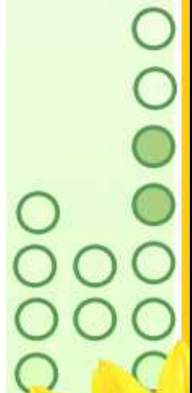
- Tyto indexy zjišťují, zda jsou ve společenstvu přítomny silně dominantní druhy nebo je společenstvo spíše vyrovnané
- Hodnoty indexů jdou opačným směrem než v případě indexů počítajících s vyrovnaností a počtem druhů (Shannon, Brillouin) a proto se často používá jejich odpočet od jedné nebo převrácená hodnota

## Simpson index

$$D = \sum \left( \frac{n_i(n_i - 1)}{N(N - 1)} \right)$$

## Berger Parker index

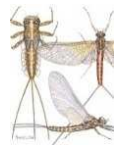
$$d = \frac{N_{\max}}{N}$$



# Berger – Parker index

## Berger-Parker index

$$d = \frac{N_{\max}}{N}$$



Species	Abundance
Baetis alpinus	736
Rhithrogena semicolorata	28
Epeorus sylvicola	8
Baetis rhodani	4
Ephemerella mucronata	1
Total number of individuals	777
Berger-Parker index	0.9472

**$N_{\max}$**  ... počet jedinců nejpočetnějšího druhu

**$N$**  ... celkový počet jedinců

Shannon diversity ↑

Berger-Parker index ↓

$1 - d$  je často použito namísto  $d$

