

Projekt 5.7 Biodiversitet - modellering af autentiske data

Indledning:

I biologi er det ofte vanskeligt i praksis at gennemføre en variabelkontrol, hvor man fokuserer på sammenhængen mellem nogle få variable, mens alle andre variable holdes konstant. Derfor benytter biologi ofte statistiske modeller, der rummer en gennemsnitsopførsel, mens afvigelserne fra modellen dels tilskrives andre variable, der ikke er under kontrol, og derfor kan have rimeligt tilfældige værdier, samt de tilfældige fluktuationer, der er et uundgåeligt træk ved statistiske modeller.

Udtrykket *biologisk diversitet* blev opfundet af biologen Thomas Lovejoy i 1980, og få år efter blev det erstattet af ordet biodiversitet, der viste sig at have stor gennemslagskraft i den offentlige debat, eller når det blev brugt af forskere og ngo'ere. Begrebet har ikke en simpel definition, men ifølge wikipedia definerer de fleste biologer

biodiversity as the "totality of genes, species, and ecosystems of a region". An advantage of this definition is that it seems to describe most circumstances and presents a unified view of the traditional three levels at which biological variety has been identified:

species diversity
ecosystem diversity
genetic diversity

In 2003 Professor Anthony Campbell defined a fourth level: Molecular Diversity.

(For yderligere information og litteratur, se: <http://en.wikipedia.org/wiki/Biodiversity>)

Under alle omstændigheder gælder det, at for biodiversitet er antallet af arter i et bestemt område den afgørende variabel. Denne variabel kaldes ofte for S (Species). Den er selvfølgelig afhængig af mange ting, men en oplagt variabel, den kan afhænge af, er arealet A af det område der undersøges. Både A og S er positive variable, så grafen for denne variabelsammenhæng ligger i første kvadrant.

Ydermere viser erfaringen, at alt andet lige så vokser antallet af arter med arealet: Jo større areal, jo flere arter er der plads til. Men antallet vokser ikke lige så hurtigt som arealet: Hvis arealet bliver dobbelt så stort (dvs. vokser med 100%), vokser antallet af arter måske kun med 50%. Det peger mod en potensmodel:

$S = c \cdot A^z$, hvor potensen z ligger mellem 0 og 1, så væksten er langsommere end for en direkte proportionalitet.

Modellen understøttes af en mangfoldighed af undersøgelser, men der er altså tale om en statistisk model med en betydelig usikkerhed indbygget. Modellen blev første gang fremsat i 1921 af den svenske biokemiker Olof Arrhenius (søn af nobelpristageren Svante Arrhenius). Ligningen kaldes derfor også for Arrhenius ligning, men i faglitteraturen er den dog mest kendt under navnet SAR-modellen (Species-Area-relation).

Den følgende tabel viser sammenhængen mellem antallet af plantearter på en ø og arealet af øen for 86 småøer i det græske øhav. Du kan hente tabellen [her](#). Her står A for øens areal (i km^2), S for antallet af arter, E for øens højde (i meter), D for afstanden til nærmeste store ø (i km) og H for antallet af habitater/biotoper, dvs. karakteristiske leveområder for planterne, fx søer, moser, enge osv.

Øens navn	A	S	E	D	H
Ag. Kyriaki	0.150	59	76	1.3	3
Ag. Nikolaos	0.119	69	30	0.4	3
Antidragonera	0.150	89	40	0.6	3
Archontonisi	0.028	11	15	0.2	1
Arefousa	0.175	43	65	0.7	3
Aspronisi (east)	0.007	11	15	2.1	2
Aspronisi (east ^o 1)	0.037	34	30	1.4	2

Aspronisi (north)	0.056	45	30	1.8	2
Aspronisi (northwest)	0.048	46	25	1.1	2
Aspronisi (west)	0.010	7	15	1.9	2
Diabates (east)	0.067	50	5	0.04	1
Diabates (west)	0.067	16	5	0.06	1
East Gourná	0.008	33	10	0.5	1
Faradonisi (northwest)	0.040	33	10	1.3	2
Faradonisi (south)	0.025	22	5	0.7	1
Faradonisi (southwest)	0.020	19	10	0.9	1
Faradonisi megalo	0.160	60	55	0.8	3
Fragkonisi	0.225	103	75	4	3
Glaronisi (north)	0.030	57	15	0.7	2
Glaronisi (south)	0.090	73	28	1.3	3
Ilias	0.023	6	10	0.03	1
Imia (east)	0.017	17	20	10.2	2
Imia (west)	0.020	20	15	9.8	2
Kalapodi megalo	0.039	55	25	3.2	3
Kalapodi mikro	0.005	12	5	3.1	1
Kalovolos	0.307	68	66	1.6	3
Kapelo	0.009	1	10	4	1
Kapparonisi	0.068	57	18	1.8	3
Katsaganaki	0.002	16	10	0.1	1
Katsagani	0.090	72	30	0.2	2
Kombi	0.090	68	20	0.4	1
Kommeno nisi	0.028	34	10	1.2	1
Koukonisi	0.472	11	10	0.2	2
Kouloura°1	0.078	76	20	0.7	2
Kouloura°2	0.020	45	30	0.2	2
Koumaro	0.100	37	20	0.1	2
Kounelonisi	0.230	59	50	1.7	3
Lidia	0.035	15	27	0.8	1
Lyra	0.050	55	40	0.2	3
Makronisi	0.034	15	30	0.4	3
Makronisi°1	0.261	76	40	0.4	3
Makronisi°2	0.197	58	30	2.1	3
Marathi	0.355	90	51	0.6	3
Mavra (east)	0.148	38	20	40.6	2
Mavra (west)	0.132	32	20	41.4	2
Megali Dragonera	0.320	109	36	0.6	3
Megalo Stroggylo	0.030	12	29	1.2	1
Megalo Trachili	0.225	11	5	0.2	2
Mikro Trachili	0.135	6	5	0.1	1
Minaronisi	0.021	45	20	0.4	2
Neronisi	0.500	27	63	0.2	3
Nisida Manoli	0.029	55	30	1	2
Paplomata	0.004	26	3	0.1	1
Patelidi	0.025	8	5	0.7	1
Piáto	0.060	56	20	1.8	1
Piganousa	0.350	101	139	0.7	3
Pitta	0.024	22	20	8	2
Plakousa	0.050	17	10	0.6	2
Plochoros	0.067	60	20	0.7	2
Pontikos	0.103	18	30	0.8	3
Prassonisi	0.013	14	10	1	1
Prassonisi°3	0.040	15	13	0.2	1
Prassonisi°1	0.011	13	2	1.3	1
Prassonisi°2	0.012	32	15	8	1
Prassou	0.500	97	40	8	4
Psathi	0.052	67	20	1.5	2
Psathonisi	0.127	44	10	0.7	2
Psonos	0.071	93	30	1.3	2
Saraki	0.007	16	30	7.8	1
Spartonisi	0.025	39	15	1.3	2
Stroggyli	0.096	62	20	1.1	2
Stroggyli°1	0.207	67	91	1.1	2
Stroggyli°2	0.150	44	76	0.5	3
Thimonies	0.010	8	10	0.1	1
Tiganaki	0.042	55	20	0.1	1
Tigani	0.140	12	5	0.03	1
Trypiti megali	0.072	72	30	1	2
Trypiti mikri	0.020	44	15	0.8	1
Vatopoula	0.007	54	15	2.4	1

Vatos	0.386	93	30	2.6	3
Velona	0.070	63	15	0.3	3
West Gourna	0.006	7	8	0.5	1
Zouka (Megali)	0.028	86	20	0.4	1
Zouka (Mikri)	0.008	79	15	0.2	1
(unnamed°1)	0.005	12	10	1.2	1
(unnamd°2)	0.0005	1	2	1.7	2

- a. Fremstil nu et diagram, der viser antallet af arter som funktion af arealet. Find den bedste potensfunktion, der beskriver denne sammenhæng. Hvad er eksponenten? Hvad er forklaringsgraden? Fremstil også et residualplot for sammenhængen. Ser det ud som om residualerne er tilfældigt fordelt?

Det kan være du blev lidt skuffet over modellen. Men husk at hvis der *ingen* sammenhæng er mellem antallet af arter og arealet, så ville en regressionslinje være vandret. En forklaringsgrad der fx kun er på 0,25 fortæller os på den ene side, at der ikke er tale om en lovmæssighed, som vi kender fra fysik, men på den anden side, at der trods alt *er tale* om en vis sammenhæng. Det er vigtigt, at man ikke nøjes med at se på regressionslinjen og forklaringsgraden, men altid inddrager residualerne.

En så lav forklaringsgrad fortæller os samtidig, at der er andre forklaringsvariable i spil. Det er ikke sikkert, vi umiddelbart kan afdække disse, men det åbner for nye undersøgelsesfelter.

- b. Undersøg nu om der med rimelighed kan påvises en potens-sammenhæng mellem øens højde E (for Elevation) og antallet af arter S , henholdsvis øens afstand D til nærmeste store ø og antallet af arter S .
- c. Konklusion: Hvilke variable synes at have indflydelse på antallet af arter? Er der en indbyrdes sammenhæng mellem de forklarende variable, som du har fundet?

Når man skal undersøge diversiteten indenfor et område, er antallet af habitater også afgørende, da forskellige habitater giver levedmuligheder for forskellige arter. Jo flere habitater der er, jo flere arter er der derfor. Denne gang er sammenhængen knap så klar, bl.a. fordi antallet af habitater i realiteten kun dækker værdierne 1, 2 og 3 (idet 4 habitater kun forekommer en eneste gang).

- d. Undersøg sammenhængen mellem antallet af habitater H og antallet af arter S . Hvad hedder den simplest mulige sammenhæng i den foreliggende situation? Konklusion?
- e. I stedet for kun at se på arealet kan man finde et bedre samlet mål for artsunderlaget ved at gange arealet med antallet af habitater. Denne variabel kaldes K og modellen, der beskriver antal arter S som en potensfunktion af K kaldes Choros-modellen, der blev fremsat i 2003.
- f. Giver Choros –modellen en bedre beskrivelse af sammenhængen end SAR-modellen?