

LP 5 Rezumat

3.8. Sinteză numerică univariată. Variabile calitative.

Tratare cantitativă și tratare specifică (calitativă).

3.8.1. Tratarea cantitativă a unei variabile calitative

- Numai după binarizare.
- **Tendința centrală** - *proporția* care ne interesează: F ori $1-F$.
- **Împrăștierea** (în jurul tendinței centrale) - în ambele situații, *dispersia* $F \cdot (1 - F)$.

3.8.2. Tratarea calitativă (specifică) a unei variabile calitative

1° Entropia informațională - măsură a diversității unei biocenoze
Entropia informațională H_p măsoară cantitatea de nedeterminare a unui experiment de capturare la întâmplare a unor exemplare din diverse specii ale unei biocenoze. Sn **informațională** deoarece "informație = nedeterminare înlăturată".

$$H_p = - \sum_{j=1}^p F_j \cdot \log_a F_j, \text{ cu } a > 1,$$

unde F_j este frecvența relativă a speciei j dintre cele p specii.

+ 2° Entropia informațională. Definiție, proprietăți

- H_p se "măsoară" în **biți** dacă $a = 2$, în **niți** dacă $a = e$ și în **diți** dacă $a = 10$.
1 bit = $1 / 1,442695$ niți = $1 / 3,321928$ diți.
- $H_p \geq 0$. $H_p = 0 \Leftrightarrow p = 1$.
- H_p maxim ($H_{p, \max}$), pentru p fixat, corespunde distribuției uniforme și are valoarea $\log_a p$.
- **1 bit** = entropia informațională a unei biocenoze cu două specii cu abundențe egale.
- Entropia informațională a unei biocenoze cu p specii cu abundențele egale crește cu numărul p de specii.

3° Entropia relativă

$$H_{rel} = \frac{H_p}{H_{p,max}} = \frac{H_p}{\log_a p}$$

- H_{rel} este adimensional.
- $0 \leq H_{rel} \leq 1$.
- $H_{rel} = 0 \Leftrightarrow p = 1$.
- $H_{rel} = 1 \Leftrightarrow$ distribuția de abundențe este uniformă.
- H_p este omologul lui σ^2 sau al lui σ , iar H_{rel} este omologul lui $CV\%$.
- H_{rel} este cu atât mai mare cu cât distribuția este mai apropiată de distribuția uniformă, regulată, echitabilă; de aceea, este o măsură a conceptului de **regularitate** sau **echitabilitate** utilizat în ecologie.
- Prin H_{rel} putem compara două biocenoze cu număr egal sau diferit de specii, în timp ce prin H_p nu putem compara decât biocenoze cu același număr de specii.

+ 5° Interpretarea ecologică a diversității specifice (Legendre & Legendre)

- Numărul de specii p depinde în mod direct de stabilitatea mediului și de capacitatea de suport a sa, am adăuga noi.
- Regularitatea distribuției de specii depinde în mod invers de activitatea biologică din mediu.

Lp 5 Teste, exerciții și probleme

TG5. Durata 200'' pe calculator.

Binarizarea variabilei calitative reprezintă:

1. studierea unei singure valori în opoziție cu ce rămâne în afara ei
2. studierea unei singure variante în opoziție cu ce rămâne în afara ei
3. studierea unei singure variante

Alegeți afirmația ERONATA:

1. vorbim de informație dacă ne aflăm înaintea experimentului și de rezultate după efectuarea lui
2. vorbim de nedeterminare înaintea experimentului și de informație după efectuarea sa

Alegeți afirmația ERONATA:

1. prin H_p comparăm biocenoze cu același număr de specii
2. prin H_{rel} putem compara două biocenoze cu număr egal sau diferit de specii
3. prin H_{rel} comparăm specii cu același număr de indivizi

Alegeți afirmația ERONATA:

1. regularitatea distribuției de specii este proporțională cu stabilitatea mediului
2. numărul de specii depinde direct de stabilitatea mediului
3. regularitatea distribuției de specii depinde în mod invers de activitatea biologică din mediu

Pentru o variabilă calitativă, tratată cantitativ, este un indicator de tendință centrală:

1. produsul $p(1-p)$, unde p este proporția unei variante
2. proporția p a unei variante

Entropia este maximă pentru distribuția:

1. normală
2. uniformă
3. gaussiană

Legendre și Legendre consideră diversitatea formată din:

1. două numere: numărul de specii și regularitatea
2. două numere: numărul de specii și entropia
3. două numere: entropia și regularitatea

Daca regularitatea unei distributii de abundente scade, numarul de specii ramanand acelasi, putem afirma ca in biocenoza respectiva:

1. a scazut activitatea biologica, foarte probabil din cauza imbogatirii resurselor din motive sezoniere, etc.
2. a scazut capacitatea de suport a mediului
3. a crescut activitatea biologica, foarte probabil din cauza diminuarii resurselor din motive sezoniere sau poluare, etc.

TC5. Durata 4'.

1. Tratarea cantitativă a unei variabile calitative presupune _____ acestei variabile, adică studierea unei singure _____ în opoziție cu celelalte.
2. Entropia informațională este o măsură a _____ unei biocenoze.
3. " _____ înseamnă nedeterminare _____."
4. Minimul de _____ dintre toate experimentele distribuite uniform este conținut de experimentul cu _____ rezultate posibile.
5. H_p este omologul _____ sau al _____, iar H_{rel} este omologul _____.
6. Conform interpretării diversității specifice, numărul de specii p va depinde _____ de capacitatea de suport a mediului, iar regularitatea distribuției de specii va depinde _____ de activitatea biologică.
7. Comparând probele 1 și 2 luate din aceeași biocenoză la momente diferite, se constată că $p_1 > p_2$, ceea ce semnifică o _____ a capacității de suport a mediului, și $R_1 > R_2$, ceea ce indică o _____ a competiției.

Exerciții sau probleme rezolvate

EGA 1.

Pentru urmatoarea distributie de abundente:

Nr.specie j	abundente absolute N_j
1	1000
2	200
3	13
4	300
5	4

sa se calculeze manual entropia, entropia maxima (in biti) si entropia relativa. Abundentele relative si rezultatele finale se vor rotunji la 2 zecimale. (Se va folosi TABELA de valori ale functiei $-F \cdot \log_2(F)$.*.) Se da $\log_2(5) = 2.3219^\dagger$ biti.

Rezolvare:

LEGENDA: \$=1 punct. 1\$ din oficiu.

Nr. specie j	abundente absolute N _j	\$abundente relative F _j =N _j /N	\$\$\$ -F*log ₂ (F) din tabel
1	1000	.66‡	.3956
2	200	.13	.3824
3	13	.01	.0664
4	300	.2	.4644
5	4	0	0
\$N=\sum N_j = 1517		\$Entropia(H5) = 1.309 biti ≈ 1.31 biti	

\$Entropia maxima (H5,max)=log₂(n)=log₂(5) = 2.3219 biti ≈ 2.32 biti

\$Entropia relativa=(H5/H5,max)=.5637625 ≈ .56

2.

Fie distribuția de abundențe absolute:

$$\begin{pmatrix} x_j \\ N_j \end{pmatrix} = \begin{pmatrix} S_1 & S_2 & S_3 & S_4 & S_5 \\ 7 & 6 & 4 & 3 & 2 \end{pmatrix}$$

Calculați entropia în biți, niți și diți, precum și entropia relativă.

Rezolvare:

$$p = 5$$

$$N = \sum N_j = 7 + 6 + 4 + 3 + 2 = 22$$

Distribuția de frecvențe relative:

* consultăm tabela din Anexa 6

† 2.3219 înseamnă 2,3219.

‡ .66 este forma sub care afișează calculatorul numărul 0,66

$$\begin{pmatrix} X_j \\ F_j \end{pmatrix} = \begin{pmatrix} S_1 & S_2 & S_3 & S_4 & S_5 \\ 7 & 6 & 4 & 3 & 2 \\ 22 & 22 & 22 & 22 & 22 \end{pmatrix} = \begin{pmatrix} S_1 & S_2 & S_3 & S_4 & S_5 \\ 0,32 & 0,27 & 0,18 & 0,14 & 0,09 \end{pmatrix}$$

Entropia în biți va fi:

$$H_5 = -\sum F_i \cdot \log_2 F_i = -0,32 \cdot \log_2 0,32 - 0,27 \cdot \log_2 0,27 - 0,18 \cdot \log_2 0,18 - 0,14 \cdot \log_2 0,14 - 0,09 \cdot \log_2 0,09 = (\text{conform tabelii}) = 0,526 + 0,51 + 0,4453 + 0,3971 + 0,3127 = 2,1911 \text{ biți} = 2,1911 / 1,442695 \text{ niți} = 1,51875 \text{ niți} = 2,1911 / 3,321928 \text{ diți} = 0,659587 \text{ diți}.$$

$$H_{5,max} = \log_2 5 = 2,3219.$$

$$H_{rel} = H_5 / H_{5,max} = 2,1911 / 2,3219 = 0,9437.$$

3.

O biocenoză a fost studiată la 3 momente de timp diferite, prelevându-se câte o probă după cum urmează în tabela cu chenar dublu:

<i>Specia</i>	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	Σ	<i>p</i>
proba 1	147	93	19	8	267	4
proba 2	160	64	15	0	239	3
proba 3	122	87	22	19	250	4

Considerând că fiecare probă reprezintă fidel compoziția biocenozei la momentul de timp respectiv:

(a) să se calculeze pentru fiecare probă entropia măsurată în biți și entropia relativă și

(b) să se precizeze cum variază în timp capacitatea de suport a mediului, precum și gradul de competiție între speciile biocenozelor. Se dau: $\log_2 4 = 2$; $\log_2 3 = 1,584962$.

Rezolvare:

(a) La tabela de frecvențe absolute de mai sus am adăugat, pentru fiecare probă, coloana de totaluri și cea cu numărul de specii, *p*. Apoi, cu ajutorul totalurilor, calculăm pentru fiecare probă frecvențele relative, entropia *H* și entropia relativă *H_{rel}*:

<i>Specia</i>	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>H</i>	<i>H_{rel}</i>
proba 1	0,55	0,35	0,07	0,03	1,4249	0,71
proba 2	0,67	0,27	0,06	0	1,1406	0,72
proba 3	0,49	0,35	0,09	0,08	1,6386	0,82

(b) Capacitatea de suport (fiind proporțională cu numărul p de specii) a scăzut de la momentul 1 la momentul 2 și a crescut de la momentul 2 la momentul 3. Gradul de competiție interspecifică (fiind invers proporțional cu H_{rel}) a rămas aproximativ același la momentul 2 față de momentul 1 și a scăzut la momentul 3.

4.

O biocenoză a fost studiată la patru momente de timp diferite, prelevându-se câte o probă ca în tabelul cu chenar dublu:

<i>Specia</i>	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>h</i>	Σ	<i>p</i>
proba 0	42	51	68	30	22	13	3	0	229	7
proba 1	86	39	49	40	21	179	6	4	424	8
proba 2	50	22	163	0	0	9	0	0	244	4
proba 3	62	31	111	35	9	120	0	0	368	6

Momentele de prelevare au fost alese după cum urmează:

Momentul 0: la scurt timp (1 an) după dezgolirea unui versant, în urma unui incendiu. (Ecosistemul este tânăr, biocenoză este alcătuită din specii pioniere.)

Momentul 1: după 20 de ani de la momentul 0. (Ecosistemul este matur, biocenoză este alcătuită din specii caracteristice zonei climatice respective, cuprinzând un strat ierbos și unul arbustiv.)

Momentul 2: după o lună de la momentul 1 când ecosistemul a fost puternic degradat prin poluare.

Momentul 3: după 10 ani de la momentul 2 când în zonă s-a desfășurat un program de reconstrucție ecologică.

(a) Să se studieze dinamica diversității biocenozei în cele 4 momente de timp.

(b) Să se interpreteze, în concepția Legendre & Legendre, starea ecosistemului în fiecare din cele 4 momente de timp. Să se indice speciile dominante și cele rare la fiecare moment de timp și să se motiveze, utilizând argumente statistice, dinamica biocenozei.

Rezolvare:

(a) Fiecare probă conține un alt număr p de specii. Deci, diversitatea lor nu poate fi comparată decât cu ajutorul entropiei relative. Pentru determinarea entropiilor relative calculăm mai întâi frecvențele relative în cadrul fiecărei distribuții de abundențe:

<i>Specia</i>	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>h</i>	Σ
---------------	----------	----------	----------	----------	----------	----------	----------	----------	----------

proba 0	0,18	0,22	0,30	0,13	0,10	0,06	0,01	0	1
proba 1	0,20	0,09	0,12	0,09	0,05	0,42	0,02	0,01	1
proba 2	0,20	0,09	0,67	0	0	0,04	0	0	1
proba 3	0,17	0,08	0,30	0,10	0,02	0,33	0	0	1

(Pentru verificarea calculului am adăugat și coloana de totaluri care va trebui să fie formată numai din valoarea 1.) Entropiile maxime se pot calcula direct pentru $p = 8$ respectiv 4 și cu ajutorul unui calculator cu logaritmi naturali ori a unui tabel de logaritmi naturali, pentru $p=7$ respectiv 6, astfel:

$$H_{7, \max} = \log_2 7 = 1,442695 \cdot \ln 7 = 1,94591 \cdot 1,442695 = 2,8074$$

$$H_{8, \max} = \log_2 8 = 3$$

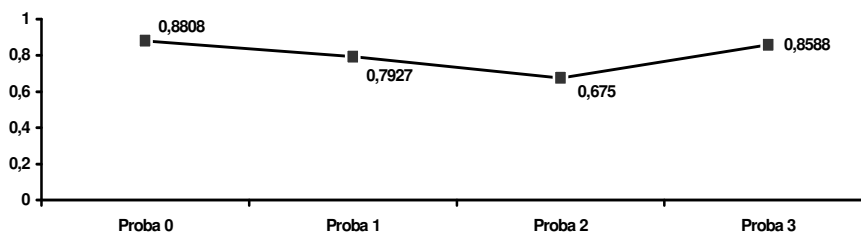
$$H_{4, \max} = \log_2 4 = 2$$

$$H_{6, \max} = \log_2 6 = 1,442695 \cdot \ln 6 = 1,79176 \cdot 1,442695 = 2,585.$$

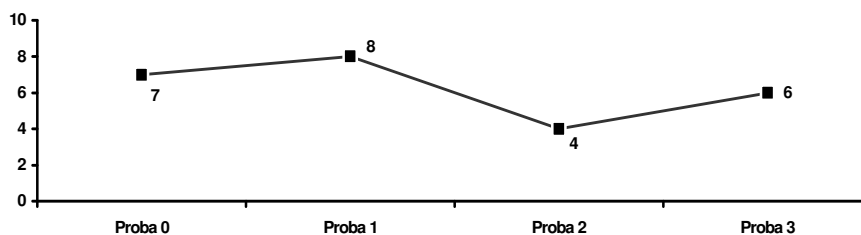
Utilizând tabela cu $-F_j \cdot \log_2 F_j$ obținem valorile de mai jos. În ultimele două coloane am adăugat entropiile H și entropiile relative H_{rel} :

Specia	a	b	c	d	e	f	g	h	H	H_{rel}
proba 0	0,4453	0,4806	0,5211	0,3826	0,3322	0,2435	0,0664	0	2,4717	0,8808
proba 1	0,4644	0,3127	0,3671	0,3127	0,2161	0,5256	0,1129	0,0664	2,3779	0,7927
proba 2	0,4644	0,3127	0,3871	0	0	0,1858	0	0	1,3500	0,6750
proba 3	0,4346	0,2915	0,5211	0,3322	0,1129	0,5278	0	0	2,2201	0,8588

Dinamica diversității (exprimată ca entropie relativă) poate fi ilustrată sub forma următorului poligon de frecvențe:



Pentru punctul următor adăugăm și reprezentarea dinamicii numărului p de specii:



(b)

Momentul:	Specii dominante:	Specii rare:
0	c, b, a	g
1	f, a	g, h
2	a, c	f
3	f, c, a	e

Urmărind cele două grafice de mai sus putem formula următoarea interpretare:
Momentul 0: Regularitatea este destul de mare (competiția este redusă) întrucât biocenoza nu este foarte complexă, ecosistemul este în stadiu tânăr.

Momentul 1: Capacitatea de suport a crescut (pentru că a crescut p) deoarece, datorită succesiunii ecologice, ecosistemul a evoluat la starea de maturitate. Din același motiv a scăzut și regularitatea (a crescut competiția). Speciile dominante sunt, mai degrabă, altele decât cele care caracterizează ecosistemul tânăr (f și a în loc de c, b și a).

Momentul 2: Din cauza poluării, capacitatea de suport a scăzut (p a scăzut dramatic), iar competiția pentru resurse a crescut, ducând la scăderea regularității. Au dispărut, în primul rând, speciile rare ale momentului anterior (g și h).

Momentul 3: După reconstrucția ecologică, capacitatea de suport a ecosistemului a fost refăcută (p a crescut), dar nu până la valoarea inițială. Competiția pentru resurse a scăzut, ducând la creșterea valorii regularității. Speciile dominante sunt o combinație între cele ce caracterizează ecosistemul matur de la momentul 1 (f și a), și cele ce caracterizează ecosistemul tânăr de la momentul 0 (c). Ecosistemul dezvoltat după reconstrucția ecologică se apropie de cel matur care a fost deteriorat, dar nu a putut totuși să fie refăcut în întregime. Speciile rare (g și h) ale ecosistemului matur, de la momentul 1, nu s-au mai restabilit după reconstrucție.

Exerciții sau probleme propuse

EGA 5.

Pentru următoarea distribuție de abundente:

Nr. specie j	abundente absolute N_j
1	34
2	56
3	76
4	3
5	2

sa se calculeze manual entropia, entropia maxima (in biti) si entropia relativa. Abundentele relative si rezultatele finale se vor rotunji la 2 zecimale. (Se va folosi TABELA de valori ale functiei $-F \cdot \log_2(F)^*$.) Se da $\log_2(5) = 2.3219$ biti.

EGA 6.

Același enunț ca în exercițiul anterior, dar pentru următoarea distribuție:

<i>Nr. specie</i>	<i>abundențe absolute</i>
1	3
2	2
3	1

EGA

Să se calculeze entropia și entropia maximă, în biți, precum și entropia relativă pentru fiecare din distribuțiile cu următoarele frecvențe absolute:

- 7.** 50 12 70. Se da $\log_2 3 = 1,585$ biți.
8. 50 50.
9. 25 25 25 25.
10. 0 100.

11.

Fie următoarea distribuție de abundente:

* din Anexa 6

<i>Specia</i> S_j	<i>Abundența</i> <i>absolută</i> N_j	Să se calculeze manual entropia H și entropia maximă H_{max} în biți, precum și entropia relativă, H_{rel} . Abundențele relative și rezultatele finale se vor rotunji la 2 zecimale. Se va folosi tabela de valori ale funcției $-F \cdot \log_2 F$. Se da $\log_2 5 \cong 2,32$ biți.
S_1	34	
S_2	56	
S_3	176	
S_4	2	
S_5	2	

12.

Fie următoarea distribuție de abundențe:

<i>Specia</i> S_j	<i>Abundența</i> <i>absolută</i> N_j	Să se calculeze manual entropia, H , și entropia maximă, H_{max} , în biți, precum și entropia relativă, H_{rel} . Abundențele relative și rezultatele finale se vor rotunji la 2 zecimale. Se va folosi tabela de valori ale funcției $-F \cdot \log_2 F$.
S_1	80	
S_2	6	
S_3	93	
S_4	142	

13.

O biocenoză a fost studiată la 3 momente de timp diferite, prelevându-se câte o probă considerată reprezentativă, dar de mărime variabilă, după cum urmează:

<i>Specia</i>	a	b	c	d
proba 1	89	149	20	6
proba 2	62	161	18	0
proba 3	85	130	21	16

(a) Să se calculeze pentru cele trei probe diversitatea specifică exprimată ca entropie măsurată în biți și regularitatea.

(b) Să se precizeze cum variază la cele 3 momente de timp capacitatea de suport a mediului, precum și gradul de competiție între speciile biocenozei la fiecare moment de timp.

Se dau: $\log_2 4 = 2$; $\log_2 3 = 1,584962$.