

## Metode statistice cu aplicații în cercetări entomologice (VIII)

Gheorghe STAN

### **Summary**

### **Statistical methods applied in entomological researches (VIII).**

In our brief synthesis at standard knowledges of beginners, the study about statistical methods in entomology approach, in this number, with the sampling programme. The following aspects are presented: the planification of the sampling and the experimental field; sampling possibilities in different of habitat units; the sampling unity and frequency of distributions; preliminary processing of the data; the pattern of the population distribution, behaviour and the type of sampling; the signification of the sampling, models of sampling and the optimum number of the samples.

### **PROGRAMUL DE ESANTIONARE**

#### *1. Modul de planificare și câmpul de lucru*

Intrucât este imposibil de a număra toate insectele dintr-un anume habitat, este necesară estimarea populației prin eșantionare (sampling, prelevare). Nu există metode universale de eșantionare iar o asemenea operațiune făcută corect la o populație de insecte trebuie analizată corelat cu distribuția și ciclul de viață al insectei. Pentru cel ce dorește să facă cercetări asupra densității și asupra nivelului real al populației unei specii, trebuie să aibă în vedere următoarele:

- să cunoască foarte bine habitatul în care va fi făcută eșantionarea (tip, întindere, delimitare, număr posibil de subunități componente definite de structura cenozei, etc);
- prin cercetări intensive și extensive, să cunoască tipul și momentul schimbărilor din populație (de ex. unele insecte dăunătoare prezintă foarte multe schimbări ale populației, de-a lungul unui sezon);
- în nici un caz, pentru estimarea abundenței populației, eroare standard să nu aibă o valoare sub o medie de 25 %;
- studiul biologiei speciei să aibă un nivel bun de acuratețe;
- eșantionarea să fie necesar corelată cu fenologia speciilor și cu fiecare model particular din evoluția populațiilor (de ex. dependent de evoluția unor dăunători sau entomofagi, cu deciziile de control sau combatere, cu diferenții factori care modifică nivelul populațiilor, evoluția stadiilor și vârstelor, evoluția vegetației și a plantei gazdă, etc);
- eșantionarea să aibă în vedere și operațiunile ulterioare de estimare a mărimiilor populației, astfel ea trebuie făcută în diferite momente și staționare; foarte multe animale și în special insectele, se pot mișca activ dintr-o parte în alta, în relație cu necesitățile de hrănire sau reproducere și în acest fel sunt afectate supraviețuirea și rata de reproducere. Mai mult, și dinamica populațiilor și aprecierea densității, depind astfel de *modelul specific al mișcărilor și deplasărilor în spațiu*. Trebuie astfel luată în considerare o terminologie adecvată care implică noțiuni ca *rata de rezidență a masculilor și femelelor, imigrarea, emigrarea, dispersia, deplasarea temporară*.
- să se pornească de la modele de eșantionare verificate și prezentate în literatura de specialitate, necesar adaptate și îmbunătățite pentru condițiile locale;

- să se lucreze cu *probe multe și mici*, decât cu probe puține și mari; este obligatoriu ca în cercetări preliminare să fie apreciat optimul de informație furnizat de un număr convenabil și rentabil de probe;

dacă studiul se face într-un singur câmp, acesta trebuie împărțit în suprafete de probă (10-20) iar dacă este necesar, și acestea să se împartă în alte subunități;

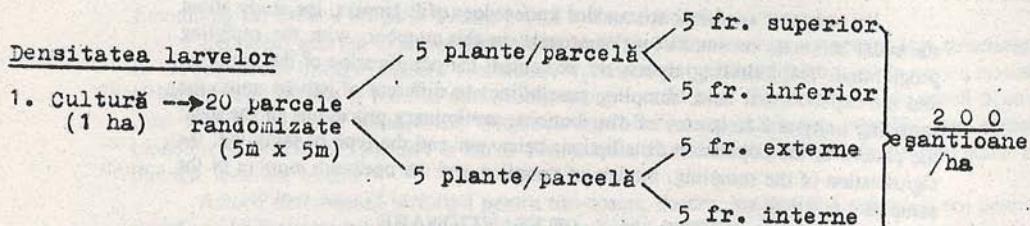
limitele de mărime ale *unității de eșantionare (proba)*, este recomandat să fie de două

*Cultură*

a. cea mai mică (ex. o frunză, o creangă) - pentru a avea un nivel cât mai mare de reproductibilitate și acesta este cazul când se lucrează cu unități multe și mici; probele mici pot prezenta dezavantajul că la densități mici se vor obține valori cu multe zerouri, situație în care interpretarea este dificilă și atunci se vor folosi probe mai mari; decizia se ia corelat cu densitatea;

b. cea mai mare și normală (adică corelat cu specia studiată).

lată ca exemplu (Fig. 31), un model simplu de planificare a unui program de eșantionare:



#### Capturare speciei adulți

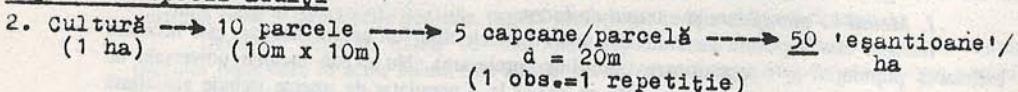


Fig. 31. Reprezentarea schematică a împărțirii a unui lot experimental în parcele și suprafete de probă, în programul de eșantionare, corelat cu scopul cercetărilor și metoda de eșantionare.

#### **2. Modalități de eșantionare în diferite unități de habitat.**

A. **Eșantionare în aer.** Dintre cele 5 tipuri de habitate, pentru insectele cu comportament de zbor, aerul este cel mai omogen și permite o multitudine de soluții. Se folosesc capcane cu suciune (aspirante), modele de capcane rotative, capcane clasice dispuse la diferite altitudini, modalități speciale de capturare (plase atașate de avioane sau planoare) sau înregistrare (radar, detectoare sensibile)(la cap. despre estimarea populațiilor - partea IX-a - va fi abordată problematica eficienței acestor metode, corelat cu mărimea insectei, viteza de zbor, viteza vântului, etc). Ca aspect general în această categorie intră toate capcanele luminoase, alimentare, feromonale sau de altă natură, care capturează insectă ca adult.

B. **Eșantionare de pe plante.** Este considerată de unii autori ca făcând parte din capitolul mare al eșantionării în mediul terestru. Acestea sunt considerate în general habitate dificile, fiind extrem de heterogene. Ca metodologie, se ia numai o parte din plantă (vârf de ramură, tulipini, frunze) constituite în unități și subunități de eșantionare (*unități de probă*), astfel că populația pe care o estimăm astfel nu este una absolută ci *este o măsură a intensității populației*. Se aplică la toate insectele adulte mai puțin mobile sau care au un comportament de dependență de aceste plante (afide, coleoptere), cât și la stadiile imature (larve, pontă). Este foarte dificil sau aproape imposibil de a elabora o metodă corectă pentru toate speciile. Există următoarele modalități de înregistrare a numărului:

a. **Numărare directă.** La unele specii este posibil pentru toți indivizii. Metoda este cea mai avantajoasă pentru că permite o estimare corectă a populației, mai ales unde există agregare mare. Unitățile de probă care se delimitizează se constituie în frunze, ramuri, tulipini sau chiar

un patrat la nivelul solului (pentru unele specii care cad imediat la atingerea plantei; aceste patrate este necesar să fie acoperite cu o plasă fină);

b. Separarea (înlăturarea) insectelor mici de pe microhabitatele pe care trăiesc:

- prin chimicale, scuturare, încălzire;
- pensulare, periere;
- spălare;
- înregistrarea urmelor (dâre, excremente, imprimări, urme de atac).

c. Colectarea indivizilor din copaci și tufișuri:

- utilizarea fileului entomologic;
- scuturare, bătaie (pe plase sau cearceafuri);
- utilizarea unor chimicale cu acțiune rapidă;
- observarea și înregistrarea descendenților (ouă, larve, pupe).

d. Extragerea insectelor din iarbă și resturi vegetale:

- instalații simple sau aparate mai complicate de sucțiune;  
- metoda de acoperire (se acoperă o anumită suprafață, se acionează cu un insecticid care în prima fază intensifică activitatea locomotoare, apoi se fac numărări);  
- sampling pentru speciile fototactice (în principiu este vorba de o cutie neagră care la partea superioară se termină cu un tub transparent, închis, în care se adună indivizi; metoda este foarte frecvent folosită pentru extragerea ipidelor din trunchiurile de molid):

- metoda Berlese-Tullgren (cu prelucrarea eșantioanelor în laborator);
- metoda capcanelor Barber.

e. Metode de eșantionare pentru speciile care pătrund în țesuturile plantelor:

- disecție;
- transparentizarea materialului biologic (cu lactofenol, NaOH 10%);
- raze X;
- diferențe între materialul curat și cel infestat (ex. dăunători la boabe);
- aspirarea indivizilor din materialul infestat;
- marcări speciale.

C. Eșantionarea insectelor parazite pe diferite gazde. Metodologia de studiu în acest caz este puțin mai complicată fiind necesare cunoștințe temeinice despre biologia și comportamentul gazdei, care poate fi nevertebrat sau vertebrat. Metodele sunt astfel individualizate de către fiecare cercetător și adaptate la grupul său.

D. Metode de eșantionare din sol și litoral. Probele se iau randomizat; pe biotopuri sau microhabitate, cu sonde speciale sau rame de diferite tipuri și mărimi. Acestea sunt cunoscute de fiecare specialist la grupul său. Noi amintim aici tehnica de sampling (vezi în continuare) și metodele generale:

- cernerea probelor uscate (sol friabil, buturugi, mușuroaie, etc);
- cernerea probelor umede (spălarea solului), spălare, flotație;
- umidificarea habitatului cu soluții diluate ce creează disconfort (dispersie, intensificarea activității locomotoare și părsirea habitatului);
- sedimentarea materialului biologic din unele probe umede;
- folosirea extractoarelor (modalitate de continuare a analizării probelor luate) *uscate* (de tip orizontal, vertical, multiple; Berlese-Tullgren, Mc Fadyen, Duffey) sau *umede*. (cu apă fierbinți, chimice, electrice).

E. Metode de eșantionare din habitate cu apă proaspătă. Se aleg în funcție de nivelul de viată al speciilor și comportamentul indivizilor:

a. Fauna de suprafață. Pentru suprafața apei sau speciile planctonice:

- plase tip plancton;
- pompe, batometre.

b. Fauna de pe vegetația din apă (metodele dă informație relativă despre densitatea populației):

- cuști de eșantionare;
- cilindrii de eșantionare;

- capcane Wisconsin și dispozitiv Gerking (vezi detalii în SOUTHWOOD, 1968).

c. Fauna de adâncime. Metodele depind de specie, natura substartului, curentul de apă, adâncime, scopul studiului, etc):

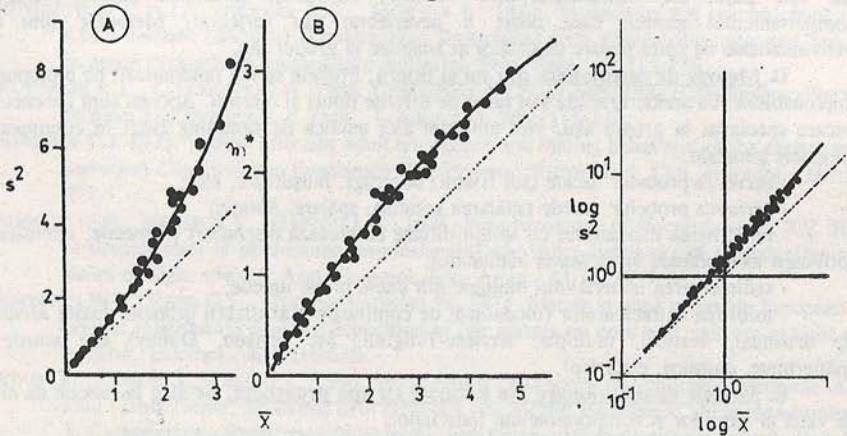
- dispozitive "plantate" care înlocuiesc porțiuni de substrat;
- dispozitive și instalații de sampling la adâncime;
- adunarea și curățarea pietrelor;
- cilindri și cuști speciale;
- aparate de dragare

### 3. Unitatea de eşantionare și frecvența distribuțiilor.

Pentru a da o perspectivă ecologică frecvenței distribuțiilor, numărul indivizilor de pe o unitate de suprafață (densitatea) este imaginat ca ceva continuu în spațiu și timp. Datele primare obținute în momentul în care se face eşantionarea (se ia o probă) se consideră date de pe o **unitate de sampling** iar indivizi din o asemenea probă constituie un întreg. Plecând de aici se vor alege apoi unitățile de probă. Unitățile ideale de sampling ecologic sunt: **linia transect**, care poate include un număr mare de habitate (aici intră de ex. și metoda capcanelor), **grilele** (permisă esantionarea la două dimensiuni și se poate face prin *numărare* sau prin semnalarea *prezență-absență*), **suprafață** (de ex. samplingul făcut pe diferite subunități - frunze, plante, porțiune de litoral) și **volumul** (de apă, sol sau aer). De obicei se evită esantionarea numai pentru o parte din indivizi și de aceea se preferă mult **probele naturale** (plante, frunze, ramuri, inflorescențe) unde *se poate face numărare integrală*. Acești procedee simplifică înregistrarea și cuantificarea dar pot să apară și confuzii, de exemplu în analiza diferențelor dintre plante sau frunze în studiul femelelor care ovipozitează.

Aprecierea și estimarea suprafeței sau volumului per individ ar fi un concept ecologic mai corect decât densitatea, dar din păcate este foarte greu de folosit. Fiecare individ ocupă un anumit volum al spațiului în care trăiește iar acesta devine tot mai mic cu cât densitatea este mai mare (ex. afidele de pe o ramură au o distribuție specifică dar este dificil de estimat suprafața ocupată).

Probele standard și frecvența de distribuție a acestora, reprezentă o metodă instantană și artificială pentru a evalua diferite tipuri de relații permanente în spațiu și timp. Trebuie să existe un număr suficient de date pentru ca esantionarea standard să poată arăta continuitatea populației cu modalitate eficientă de prelucrare statistică (Fig. 32).



**Fig. 32.** Vizualizarea reprezentării grafice a relației dintre varianță ( $s^2$ ) și media ( $\bar{x}$ ), în cazul A și dintre media grupată ( $m^*$ ) și medie ( $\bar{x}$ ), în cazul B, prin reprezentare la scară aritmetică, pe ambele axe. A. o curbă concavă descrisă de relația  $s^2 = 1,036\bar{x} + 0,323\bar{x}^2 - 0,037\bar{x}^3$ . B. o curbă convexă descrisă de relația  $m^* = 0,036 + 1,323\bar{x} - 0,037\bar{x}^2$ . C. o relație liniară, între varianță și medie, reprezentate la scară logaritmică și descrisă de formula  $\log s^2 = \log 1,35 + 1,18 \log \bar{x}$  (după TAYLOR et al. 1978).

Inițial au fost stabilite următoarele criterii pentru unitatea de sampling (unitate de probă) (MORRIS 1955):

- toate unitățile habitatului să aibă șansă egală de selectare;
- să existe stabilitate (dacă sunt schimbări, să fie minore și continuu măsurate);
- să prezinte particularități de trecere spre unitatea de areal a zonei investigate;
- folosind unitatea de probă ca și habitatul reprezentativ, proporția populației de insecte trebuie să rămână constantă;
- să fie ușor de delimitat în câmp;
- mărimea să fie acceptabilă, astfel ca raportul dintre varianță și cost sau efort, să fie rezonabil;
- dacă se începe cu o metodă de sampling, ea să fie folosită peste tot.

Intr-un sampling preliminar este necesar să se estimeze varianța fiecărei unități de sampling ( $s^2_u$ ) în așa fel ca la comparare, dacă diferențele sunt nesemnificative, să se ia ca mărime, *cea mai mică unitate de sampling* (ex. dacă cea mai mică unitate de probă la înregistrarea larvelor de *M. brassicae* de pe varză a fost de 3 verze / 1m rând, atunci varianța pentru 2m de rând s-a împărțit la 2, pentru 4m, la 4 și pentru 10m rând, s-a împărțit la 10; în acest fel costurile au devenit similare prin reducere la o bază comună -  $C_u$  - ; costul relativ net pentru fiecare unitate de probă va fi proporțional cu  $C_u S^2_u$ ; în ultima relație  $C_u = \text{costul per unitate, la o bază comună}$ ,  $s^2_u = \text{varianța per unitate}$ . Precizia relativă netă va fi:  $P_n = 1 / (C_u \cdot s^2_u)$ ).

#### **4. Prelucrare și interpretare preliminară. Studii statistice**

Modul de prelucrare și interpretare are la bază date obținute din studii similare sau cercetări personale preliminare. Despre unele aspecte am amintit anterior.

**A. Numărul de probe, sirul de valori, eliminarea valorilor aberante, normalizarea datelor** (Bul.inf. 4(2): 99-101, 1993).

**B. Transformarea datelor** (Bul.inf. 4(2): 102-111, 1993).

**C. Folosirea unor posibilități ajutătoare:** hârtia logaritmică, hârtia cu probabilități.

a. **hârtia logaritmică.** Are o liniatură specială, fie numai la nivelul axei OY fie pentru ambele axe. Se folosește în reprezentarea grafică pentru vizualizarea datelor atunci când există diferențe mari și foarte mari între valorile sirului de date. Scara logaritmică poate fi folosită numai pe una din axe (Fig. 33) sau pe ambele axe, pornind cu notarea în funcție de valorile avute. Se folosește frecvent atunci când avem valori și foarte mici și foarte mari. Aparent pare a avea unele inconveniente prin faptul că nu mai însăștează un trend semnificativ de evoluție al unei curbe care reprezintă dinamica, datorită faptului că valorile mici sunt ridicate iar cele mari coborâte.

b. **hârtia cu probabilități.** Este un mijloc eficient de scurtare a perioadei de investigare în stadiile preliminare ale analizei statistice a datelor. Folosirea ei dă posibilitatea de reprezentare grafică a normalității datelor. Estimările făcute pe acest tip de hârtie (Fig. 34) nu prezintă o mare acuratețe (pentru că este totuși dificil de a trage după ochi liniile) ca și în cazul celei calculate. Există două modalități de folosire a acestui tip de hârtie:

\* la un număr mare de probe (peste 20): ele pot fi grupate în clase de frecvență (Bul.inf., 4(2): 85-88, 1993) iar frecvențele cumulate sunt marcate pe grafic.

Exemplu: pentru a înregistra variabilitatea numărului de ouă depus de o femelă de *M. brassicae*, s-au observat 100 plante. Intervalul a fost: cel mai mic număr de ouă/frunză = 5; cel mai mare număr de ouă/frunză = 85. La hârtia cu probabilități, având scara aritmetică pe verticală, se poate marca ușor intervalul 5-85. Cea mai mică probă (5 ouă) s-a înregistrat numai pe o frunză (deci 1%, pentru că au fost 100 observări); deci pe grafic avem 1% cu 5. Următoarele situații au fost: pe două frunze s-au înregistrat 10, 13 și 15 ouă; în această situație, următoarele puncte vor fi: 2% cu 10, 3% cu 13 și 5% cu 15. Procedeul se va continua până la ultima valoare.

\*\* la un număr mic de probe: aici probele sunt aranjate (în rând) *dar nu grupate*. Punctele corespunzătoare mărimii lor sunt plasate la intervale egale de-a lungul scării de frecvențe. Cea mai mică probă este pusă la valoarea procentuală corespunzătoare la 50 / N (N = numărul de probe) iar următoarele probe sunt inserate la intervale de 100 / N% (Tabelul 16).

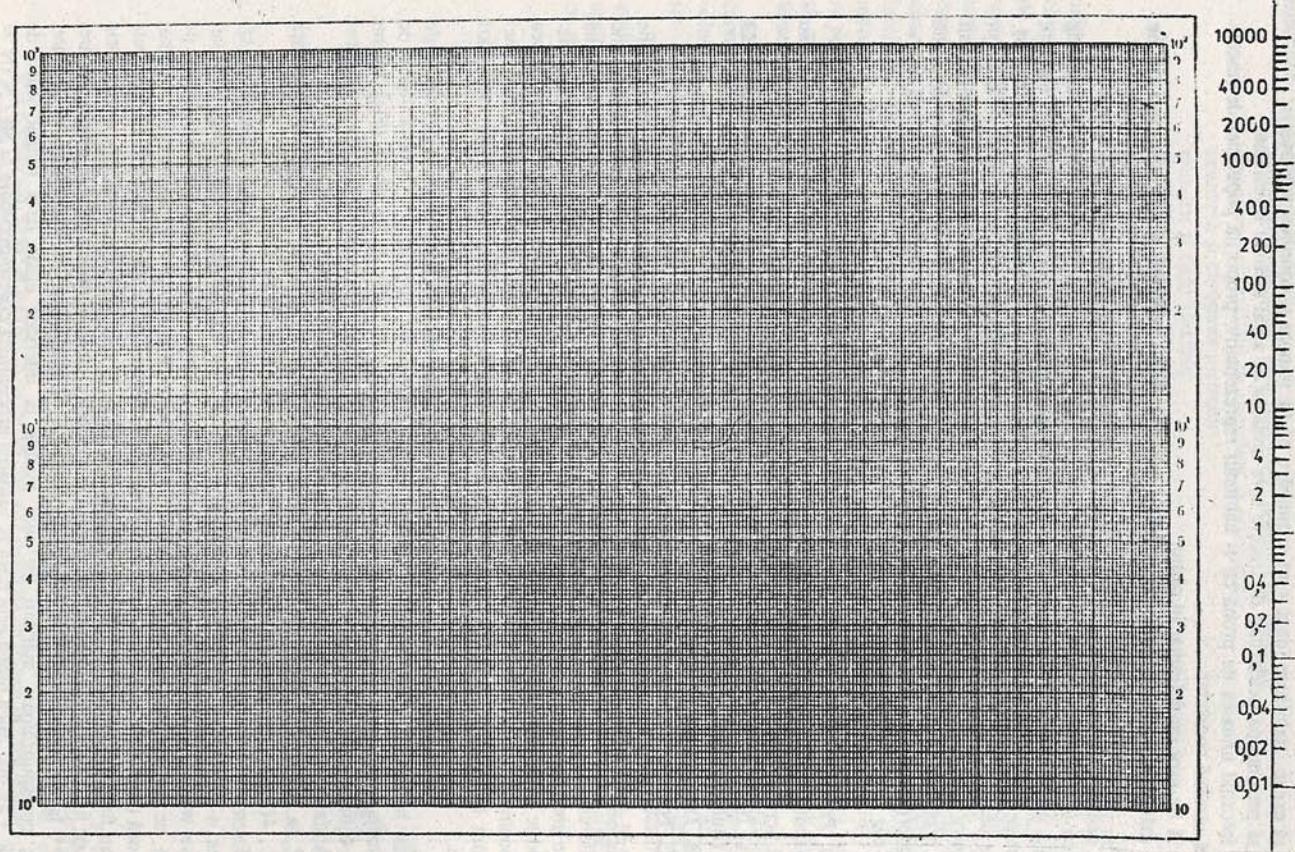


Fig. 33. Model de hârtie logaritmică utilizată în prelucrarea preliminară a unor date de eșantionaj

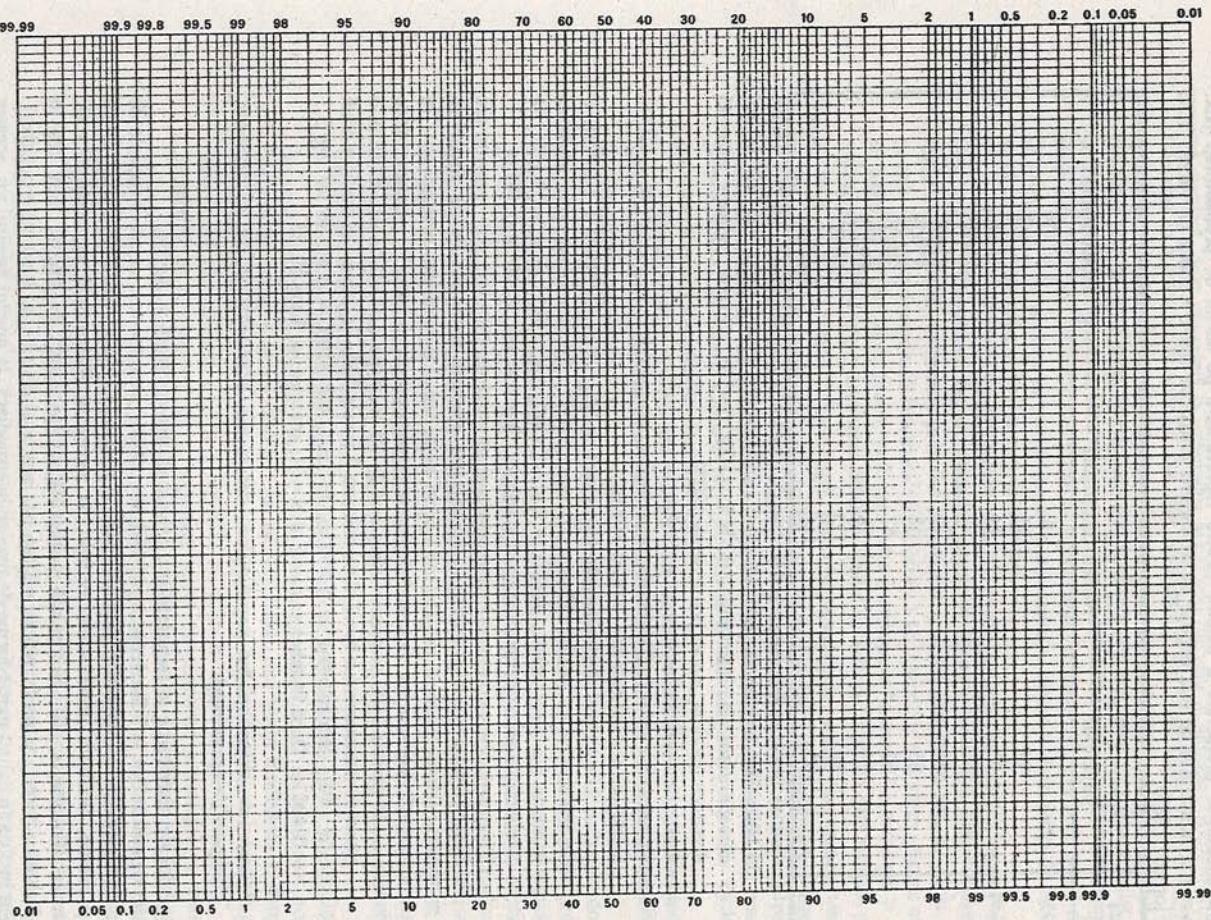


Fig. 34. Modelul hârtiei de probabilități folosite în analiza statistică preliminară a datelor de eșantionaj

Tabelul 16

Modalitate de prelucrare a datelor pentru reprezentare pe hârtia de probabilități (Ex. Variația numărului de ouă depuse de *Mamestra brassicae* într-o cultură de varză; SDE Florești, 1989)

Planta (N)	Valoarea % a punctului*	Număr ouă/plantă**	log Nr.ouă/plantă***
1	5	8	0.90309
2	15	12	1.07918
3	25	14	1.14613
4	35	17	1.23045
5	45	26	1.41497
6	55	38	1.57978
7	65	62	1.79239
8	75	78	1.89209
9	85	111	2.04532
10	95	168	2.22531

la cea mai mică probă, prima cifră este 5, deoarece corespunde la  $50/N$  ( $N=10$ ;  $50/10=5$ ); în situația dată în tabel, următoarele valori cresc din 10 în 10, corespunzător la intervalul  $100/N\%$  ( $100/10=10$ ). Ultima valoare este 95 deoarece, evident, nu se poate depăși 100%;

\* valorile sunt aranjate ordonat în sir crescător;

\*\* se calculează conform Tabelelor logaritmice și metodei de calcul descrise în lucrări de specialitate (ex. ROGAI 1983).

Logaritmul mediei numărului de ouă =  $\Sigma N_i / N = (0.90309 + 1.07918 + 1.14613 + \dots + 2.22531) / 10 = 1.53087$ . Se calculează apoi abaterea standard (s) (vezi relația /14/ din Bul.inf. 4(2):94, 1993) și  $s = 3,100$ . Se calculează apoi logaritmul abaterii standard -  $\log s = \log 3,100 = 0,49136$ .

Spre readucere amintesc, iată ca exemplu și un mod de calculare. Un logaritmul are caracteristica (partea întreagă) și mantisa (partea zecimală). Există logaritmul natural sau "în bază 10" (ln sau uneori notat  $\log_{10}$ ) și logaritmul zecimal sau "în bază 2" (log sau lg). Pentru 3.1 - caracteristica este 0; pentru 31 - caracteristica este 1; pentru 3100 - caracteristica este 3. Să luăm pe 3100. Mantisa acestuia o găsim la intersecția lui  $N=310$  cu coloana 0 și este 136. Se adaugă și primele două cifre (49), deci mantisa va fi 49136. În final,  $\log 3100 = 3,49136$  iar  $\log 3,1 = 0,49136$ . Dacă în tabel se căută la  $N=31$ , mantisa va fi tot 49136.

### 5. Modele de distribuție a populațiilor, descrierea experiențelor de eșantionare, tipuri de eșantionaj.

#### Modelul de distribuție al populațiilor. Relația dintre varianță și medie.

(Alte detalii în Bul.inf. Soc.lepid.rom. 4(2):93, 1993). Deși teoretic există 3 modele de distribuție (în agregate, întâmplătoare și uniformă), în practică, prin prelucrarea relației dintre varianță și medie, întâlnim următoarele situații: în agregate (binomială negativă) ( $s^2 > \bar{x}$ ), întâmplătoare (Poisson) ( $s^2 = \bar{x}$ ), binomială pozitivă ( $s^2 < \bar{x}$ ) (ca formă de trecere între cea în agregate și cea uniformă) și uniformă ( $s^2 = 0$ ) (un caz aparte al distribuției binomiale pozitive). Este evident însă că în natură există forme de trecere intermediară, la nivelul fiecărui tip, acestea putând fi surprinse prin reprezentare grafică. Pe de altă parte însă, o reprezentare impreciosă poate fi datorată unui studiu incorect. Distribuția spațială este una dintre cele mai caracteristice proprietăți ecologice ale speciilor. Spre deosebire de rata de creștere și reproducere, care adesea variază mai mult între generațiile unei specii decât între specii, distribuția spațială este aceea care furnizează parametrii caracteristici care deosebesc speciile. Acești parametrii sunt expresia

*populației și a comportamentului individual*. Repartiția spațială a populațiilor de insecte este foarte rar regulată, doar căteodată la întâmplare și cel mai frecvent în agregate (BADENHAUSSER 1989). Interesul în a caracteriza speciile este atât fundamental cât și aplicativ. Pentru entomologii care aplică datele trebuie avute în vedere două surse potențiale de eroare: a) unele specii de insecte se reproduc atât de rapid încât densitatea populației se poate schimba semnificativ chiar pe perioada de experimentare, se crează modificări temporale în densitatea populației datorită efectelor spațiale asupra distribuției; b) dacă datele statistice provin dintr-un asemenea câmp de probă ca cel de sus, în situația în care acestea se extrapolează (la alte culturi, câmpuri, subunități de probă), automat se obțin date statistice bizare asupra distribuției.

Pentru caracterizarea structurilor spațiale există *legile probabilității*: legea Poisson (pentru repartitia la întâmplare), legea binomial negativă (pentru descrierea distribuțiilor aggregative; legea este definită matematic prin doi parametrii - media aritmetică ( $m$ ,  $\bar{x}$ ,  $\mu$ ) și un exponent pozitiv ( $k$ ) care depinde de facultatea speciilor de a se reproduce și a cărui valoare este în funcție de numărul de unități de probă; unii autori dau o explicație "naturală" a varianței > sau < decât media, plecând de la presupunerea cănsăși media variază iar expresia care include pe  $k$  se scrie:  $s^2 = x + x^2/k$ ), legea Polya-Aeppli, legea Neyman (sunt modele care caracterizează repartitia de tip aggregativ iar aggregatele sunt acelea care au o distribuție la întâmplare), legea lui Taylor (sau relația dintre varianță -  $s^2$  - și medie -  $\bar{x}$  - și estimatează numărul de indivizi per unitatea de eșantionaj aleasă la întâmplare), metode bazate pe eșantionajul sistematic (de mare importanță pentru analiza structurii spațiale prin descrierea și interpretarea proceselor biologice și prin punerea la punct a tehniciilor de eșantionare).

Caracterizarea structurilor spațiale prin legi de probabilitate permite o reducere și sinteză a celor mai utile informații prin care se compară populațiile iar parametrii acestora se urmăresc în funcție de timp și loc.

a. **Comportamentul și distribuția randomizată**. Există distribuție randomizată? Cu excepția animalelor planctonice și cele cu dispersie aeriană, se pare că este puțin probabil să existe o distribuție randomizată (la întâmplare). În ultima categorie de mai sus intră insectele în "stadiul de zburător", fiind cele mai randomizat distribuite și utilizează intens mediul aerian. În concordanță cu necesitățile lor comportamentale, neavând acces la acesta în stadiul de larvă sau chiar ca adult bătrân când comportamentul de zbor este foarte redus. Astfel, la unele specii (ex. afidele) se realizează în realitate o puternică agregare ( $s^2 > \bar{x}$ ) dar distribuția poate deveni randomizată (Fig. 35) în cazul unei migrații puternice. La specii de insecte cu dependență mare de planta gazdă, mecanismul de dispersie poate crea aproape o distribuție randomizată ( $s^2 = \bar{x}$ ), dar aceasta este asociată cu o excesivă reproducere și o mare mortalitate.

După un comportament locomotor activ și sofisticat de găsire a gazdelor, în final, pe acestea, cercetătorul va înregistra tot o distribuție nerandomizată. În fazele cele mai sedentare ale ciclului de viață, plus speciile la care dispersia este mai puțin randomizată, distribuția nu este niciodată cu adevarat randomizată. Putem spune că toți indivizii sunt o parte a populației și ei trebuie să rămână așa pentru a supraviețui și a se dezvolta. La "marginile" unei populații densitatea se diminuează pe o suprafață agală cu maximul suprafeței pe care o cercetează individul, la acest nivel fiind limita care definește populația. Dincolo de limitele acestei densități, indivizii sunt îndepărtați, atât de populație cât și unul de altul și vor avea cu adevarat o distribuție randomizată. Iată deci semnificația cunoașterii comportamentului speciei și de aici, atenția unde trebuie făcută eșantionarea.

Legat de sampling mai poate apărea un aspect. În unele probe mici putem nimeri doar câte un individ și asta pe mai multe unități de probă. Automat aceasta implică prezența unei randomizări și nu se mai poate face distincție între adevarata și falsa randomizare. Entomologul bine pregătit știe însă că în natură o populație acoperă o anumită suprafață, dar aceasta nu este continuuă ci are "virfuri" și "goluri" care sunt în mișcare și prin necunoaștere și eșantionare insuficientă, apar distribuțiile fals-randomizate chiar la specii cu distribuție tipică în agregate (Fig. 36). În asemenea situații, mărimea probei se confundă cu densitatea, iar un singur individ este evident că nu se poate agrega. Să atunci? De obicei se intensifică eșantionarea și se încearcă să se evite asemenea probe. Printr-o eșantionare largă și extinsă se pot obține modele de distribuție

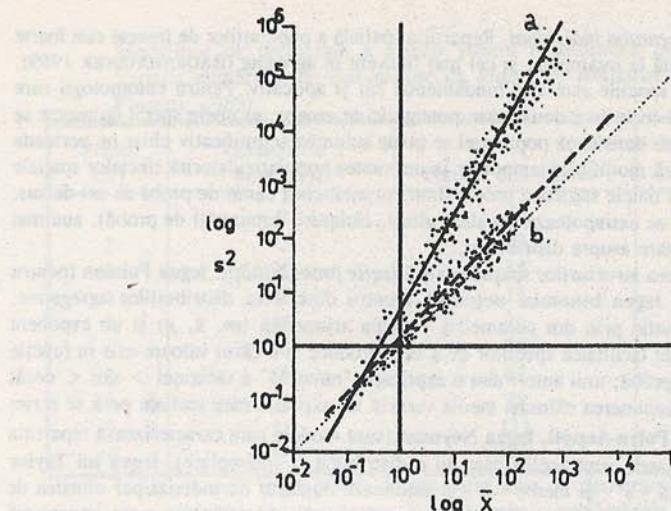


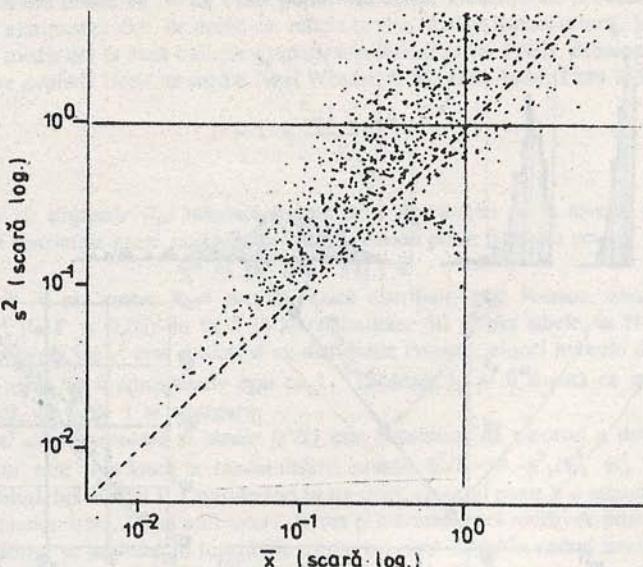
Fig. 35. Reprezentarea relației dintre medie și varianță poate surprinde două modele de distribuție, pentru o aceeași specie și același loc. Figura este preluată după TAYLOR (1974) și se referă la o populație de afide, care evidențiază o distribuție în agregate ( $s^2 > \bar{x}$ ) (a), atunci când studiul s-a făcut în cultură și o distribuție apropiată de cea întâmplătoare ( $s^2 = \bar{x}$ ) (b), atunci când s-a făcut o eșantionare aeriană, la o distanță nu mai mare de 10m de cazul anterior.

reală, proprii, repetabile, dar este absolut necesară cunoașterea comportamentului (parte esențială a distribuției spațiale), selecțarea habitatelor caracteristice și un model de sampling care se asigure o informație optimă. Fără a intra în amănunte, arăt doar că unii autori (cit. în TAYLOR 1984) disting 5 tipuri ale modelului spațial: *vectorial* (cauzat de evoluția liniară a factorilor de mediu), *reproducător* (care determină influențe asupra densității), *social* (cauzat de ansamblul comportamentului), *coactiv* (determinat de comportamentul dintre specii), *stochastic* (rezultat al tuturor proceselor întâmplătoare).

**b. Distribuția în agregate.** Am arătat în studiile noastre anterioare că în analiza distribuției, ceea ce poate fi efectiv măsurat sunt media ( $\bar{x}$ ) și varianța ( $s^2$ ). Se presupune că acestea vor fi conforme cu orice model de divizare inițială, pe culturi, pe ani, pe subunități de diferite mărimi și justifică că densitatea populației evoluează în funcție de specie, timp și spațiu. Media și varianța se pot combina în diferite moduri, pentru a ilustra coeficienții sau indicii de agregare (parametrii de dispersie). (Agregarea este folosită în înțelesul de a descrie condiția spațială în care densitatea este mai condensată decât randomizarea și ea nu are implicații de cauzalitate. O populație cu adevărat randomizată necesită o totală independență biologică și deci nici un individ al unei aceleiași specii nu poate fi influențat de către altul. Acest lucru, practic în lumea vie nu se realizează. Am arătat însă că apare falsă-randomizare datorită informațiilor insuficiente și neadecvate).

Mai mulți autori prezintă câteva atribute ale unui asemenea coeficient sau "indice perfect" care să evidențieze gradul de ne-randomizare: 1. poate prevedea valorile reale și continui pe un interval complet al modelului de distribuție; 2. poate fi influențat de numărul de unități de probă, de mărimea unităților și de numărul indivizi; 3. să poată fi ușor de calculat (GREEN 1966). Pentru un indice al distribuției spațiale, care să oscileze între *total-regulat* ( $s^2 < \bar{x}$ ) și *total agregat* ( $s^2 > \bar{x}$ ), deci cu "simetrie" față de distribuția randomizată ( $s^2 = \bar{x}$ ), LEFKOWITCH (1966) mai adaugă două atribute: 4. corespunzător cu previziunile teoretice, valorile indicelui pot fi centrale; 5. permite accesibilitatea diferențelor teste de semnificație între diferențele valorii obținute. TAYLOR

(1984) mai adaugă încă un aspect: 6). funcția descriptivă a acestor indici poate fi clar distinctă în aplicare și interpretare comparativ cu o elaborare teoretică.



**Fig. 36.** Model de reprezentare al distribuției unei populații, în condițiile în care s-au luat probe prea mici, ceea ce a generat un model artificial al relației dintre varianță și medie. S-a obținut astfel o distribuție fals-randomizată ( $s^2 = m$ ) și una fals-regulată, la nivelul mediei egal cu 1 ( $x = 10^0$ ). Acest model ilustrat aici (după TAYLOR & WOIWOD 1982) este dificil de interpretat în programele de monitoring și combatere.

Prima dintre aceste cerințe, presupune continuitatea agregării pentru toate distribuțiile statistice, deoarece un model de distribuție poate să se schimbe instantaneu în funcție de modificarea densității (Fig. 37) iar plasarea punctelor pe aceeași linie, în cadrul relației dintre varianță și medie, indică corectitudinea metodei de lucru. A doua cerință este mai contraversată. Rămâne de văzut, în cadrul fiecărei specii, dacă un indice este sau nu, influențat de numărul indivizilor, atunci când însuși comportamentul de agregare este modificat chiar de agregare.

In literatura de specialitate sunt tratați mai mulți indici de agregare. Deoarece cauzele biologice ale agregării sunt de natură comportamentală, modul lor de acțiune este imprevizibil, astfel că un asemenea mecanism biologic corespunde unei distribuții. Indicele de agregare este mai ușor de analizat decât frecvența de distribuție. Lucrările evidențiază următorii indici: parametrul "k", parametrul "b" din relația  $T = a + b \cdot r^{\alpha}$  Power Law, indicele de dispersie testat prin  $\chi^2$ , raportul  $s^2/x$ , factorul binomial negativ (k), indicele Morisita ( $I_{DM}$ ), media grupată (m), relația dintre k,  $I_{DM}$  și m, ecuația lui Breder, coeficientul de aglomerare al lui Denev, alții indici (Greene, Leffko, Vitche).

- *Parametrul "k"* (alte date sunt prezentate și în continuare) este un indice al agregării în populație. Distribuția binomial negativă poate fi dată de valorile lui k ca o măsură a dispersiei (valorii mai mici ale lui k indică o creștere a agregării, în timp ce valorii mari, peste 8, indică faptul că distribuția se apropie de una Poisson. Din păcate valorile lui k sunt influențate de mărimea probei și atunci, pentru comparare, este necesar să se păstreze aceeași unitate de mărime. Agregarea, reprezentată de distribuția binomial negativă poate fi evidențiată uneori ca o *aggregare activă*, cauzată de boli, dușmani, microclimat, plante gazdă, calitatea solului, etc. Există unele relații matematice care pot pune în evidență chiar și cauza agregării. Numărul mediu de indivizi dintr-o agregare, poate fi calculat prin relația:  $\lambda = (\bar{x} / 2k) \cdot r$  ( $\bar{x}$  = media;  $r$  =

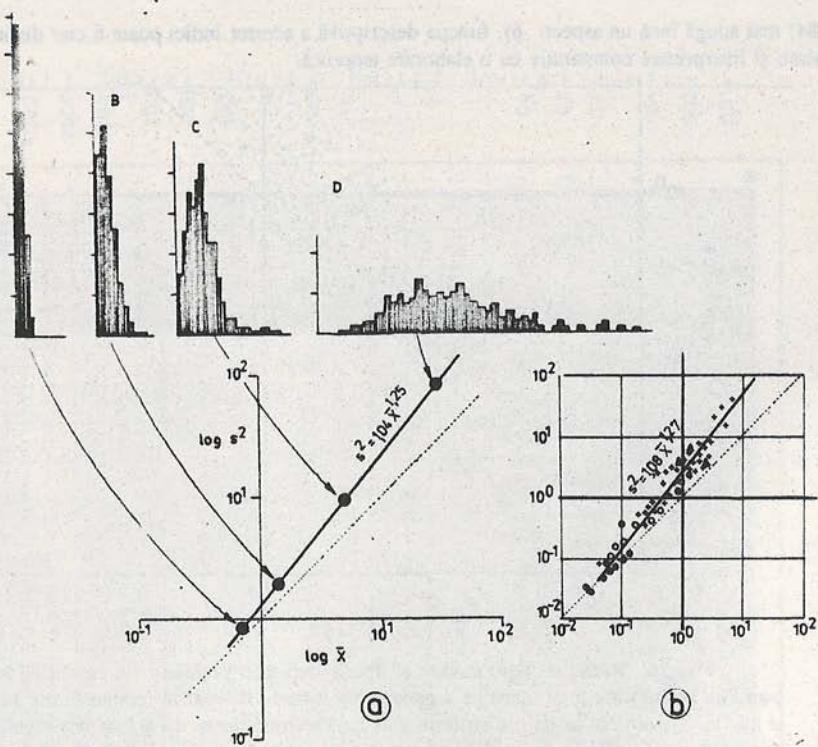


Fig. 37. Schimbarea instantanea a modelului de distribuție, de la una binomială Poisson (A) la densități mici, prin distribuția binomial negativă (B) și una neajustată (C), la densități medii, până la tipul Neyman (D) la densități mari (date din McGUIRE et al. 1957, pentru *Ostrinia nubilalis*). Toate cele 4 distribuții sunt conforme cu relația  $s^2 = 1,04\bar{x}^{1.25}$ , date preluate din TAYLOR (1965) (a). Similar, la *Lymantria dispar* distribuția ouălor se schimbă odată cu creșterea densității de la una Poisson (cerc plin), la una logaritmică naturală (triunghi) trecând prin una Poisson și binomial negativă (cerc gol) și una numai binomial negativă (cruce), după relația  $s^2 = 1,08\bar{x}^{1.27}$  (b) după datele din BROWN & CAMERON (1982). Graficele sunt preluate din TAYLOR (1984).

este o funcție cu distribuție  $\chi^2$  cu  $2k$  grade de libertate;  $\lambda = \text{nr. de indivizi dintr-o agregare pentru un anumit nivel de probabilitate dat lui } r$ . Probele pot fi luate din diferite câmpuri sau habitate și fiecare va avea un  $k$  separat. Compararea acestora și calcularea unui  $k$  comun este necesară în transformarea datelor pentru analiza varianței și pentru *samplingul secvențial*. În acest caz se poate folosi metoda regresiei (BLISS & OWEN 1958). Apoi, pentru fiecare unitate, pentru reprezentare grafică, se calculează următorii doi parametrii statistici:

$$x' = (\bar{x}^2 - s^2) / N ; \quad y' = s^2 - \bar{x}$$

( $\bar{x} = \text{media}; s^2 = \text{varianța}; N = \text{numărul de indivizi numărați, pe baza cărora se bazează estimarea mediei}.$ )

Pentru grafic, pe cele două axe se reprezintă valorile  $x'$  și  $y'$ , acestea calculându-se pentru

varianta concretă de lucru.

- *Parametrul "b"*. În cadrul modelelor de descriere a experiențelor de sampling, relația varianță - medie ocupă o atenție deosebită. Despre Legea Taylor's Power Law am vorbit pe larg. Deși exponentul  $b$  din relație ( $s^2 = \bar{x}^b$ ) este puțin, sau deloc, influențat de schimbarea locului de eșantionaj, de mărimea probei, de pesticide, relația în sine rămâne uneori critică. Relația I<sub>WAO</sub> dintre varianță și medie are la bază indicele grupării mediilor (LLOYD 1960), denumit și **indicele Lloyd** (L) care depinde liniar de medie (vezi WOLDA & MAREK, 1984; LEPS 1993):

$$L = 1 + \frac{s^2 - \bar{x}}{\bar{x}^2}$$

- *Indicele de dispersie* ( $I_D$ ) măsoară îndepărtarea distribuției de la nivelul randomizat. Măsura în care o distribuție poate satisface un model Poisson poate fi testată prin  $\chi^2$  în relația:

$$\chi^2 = [s^2(N - 1)] / \bar{x}$$

( $s^2$  = varianță;  $N$  = nr. probe;  $\bar{x}$  = media). Dacă distribuția este Poisson, atunci valoarea calculată a lui  $\chi^2$  (la  $P = 0,05$ ) nu va fi în afara limitelor lui  $\chi^2$  din tabele, la  $N-1$  grade de libertate. Dacă valoarea lui  $\chi^2$  este conformă cu distribuția Poisson, atunci indicele de dispersie,  $I_D = \chi^2 / (N - 1)$  va fi aproximativ egal cu 1. O valoare  $I_D = 0$  denotă că indivizi sunt uniform distribuiți, iar  $I_D > 1$  = agregare.

- *Raportul dintre varianță și medie* ( $s^2/\bar{x}$ ) este considerat ca raportul a două variante. Deoarece raportul este o măsură a randomizării, atunci  $s^2/\bar{x} = s_1^2/s_1^2$ , ( $s_1^2$  = varianță întâmplătoare echivalentă mediei  $x$ ). Considerând în acest fel, raportul poate fi o măsură a deviației de la distribuția randomizată. După unii autori, forma și mărimea unei unități de probă, afectează acest raport. Raportul se schimbă în funcție de medie ( $x$ ). Într-un fel în cadrul unei specii și în alt fel între specii (Fig. 38). După alii autori raportul este un "inducător în eroare cinstit", ca măsură a agregării.

- *Indicele Morisita* ( $I_{DM}$ ) este o măsură a modelului de dispersie și a fost conceput ca, independent de frecvența distribuțiilor, pentru a înregistra efectul mărimii unei probe (patrat) asupra mediei numărului de indivizi (când patratul este suficient de mic, distribuția devine randomizată!). Acest parametru introdus într-o ecuație cu indicele de diversitate Simpson, este teoretic independent de numărul total, media densității, mărimea și numărul de patrate, considerând că fiecare probă este luată randomizat dintr-un număr de  $q$  grupuri ale unei populații infinite ( $q$  este numărul de patrate). Folosind acest indice, NAKASUJI et al. (1965) au analizat distribuția spațială a adulților și pontei, în *relație cu comportamentul*, la 3 specii de insecte. Concluzia a fost că habitatele pentru depunerea pontei și comportamentul, au determinat distribuția pontelor. Iată mai jos două relații ale acestui indice:

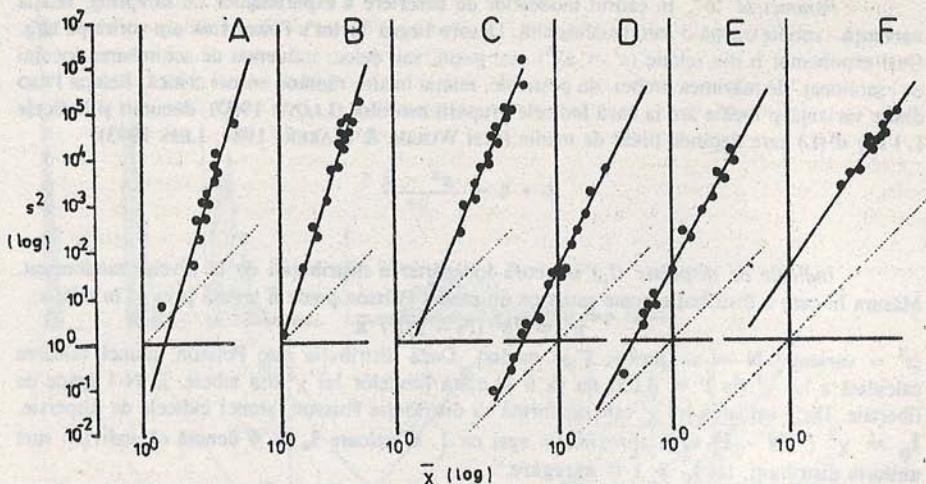
$$I_{DM} = N \cdot \frac{\sum n_i(n_i - 1)}{\sum x(\sum x - 1)}$$

( $N$  = nr. total de probe;  $n_i$  = nr. indivizi/probă;  $\Sigma x$  = suma nr. indivizi din toate probele)

$$I_{DM} = I_{dm} \cdot \frac{\frac{(G - 1)}{N}}{\frac{G}{G - 1}}$$

( $G$  = nr. subunități, perioade, grupe, zile, ore, etc;  $N$  = nr. total de probe (eșantioane) din toate seriile;  $I_{dm}$  = indicele de diversitate calculat separat pentru fiecare grup).

Acest indice poate compara distribuția indivizilor în cadrul perioadelor din zi sau pentru subunitățile de probă ale unei suprafețe mai mari (ex. se compară nivelul de captură din capcane din primele 6 zile, cu nivelul din următoarele 6 sau mai multe zile; distribuția larvelor pe frunze în diferite puncte ale coroanei unui copac, distribuția adulților la nivelul diferitelor inflorescențe,



**Fig. 38.** Varianța probelor din câmp poate crește cam la aceeași rată ca și densitatea medie (F) ori ca patratul mediei (D) sau cubul mediei (B) sau cu puterea a 4-a a mediei (A). Fiecare punct reprezintă o probă la nivel de an (vezi D) în diferite puncte de eșantionare, pe o suprafață mare. Varianța poate fi acum cu acuratețe preisă, la o densitate dată a mediei. Datele din această figură se referă la adulți de: A - *Aphis sambuci* ( $b=3,39$ ), B - *Lycophotia porphyrea* ( $b=3,11$ ), C - *Agrotis exclamacionis* ( $b=2,64$ ), D - *Euproctis chrysorrhoea* ( $b=2,04$ ), E - *Cymatophorima diluta* ( $b=1,96$ ), F- *Epirrhoe alternata* ( $b=1,37$ ) (după TAYLOR et al. 1980).

- *Media grupată* ( $m^*$ ) este un indice care reprezintă încercarea de a lega distribuția spațială de comportamentul dependent de densitate (LLOYD 1967), concept care este considerat celalalt față a "mediei distanței până la cel mai apropiat vecin" (CLARK & EVANS, 1954; vezi în continuare la estimarea mărimei populației).  $m$  este suma densităților medii, implicând raportul varianță-medie și reprezentând efectul de aglomerare, mai puțin individul din centru. Se scrie:

$$m^* = \bar{x} + (s^2/\bar{x} - 1)$$

- *Relația dintre k, I<sub>DM</sub> și m*, este următoarea:

$$1/k + 1 = I = m/x$$

- *Ecuația lui B r e d e r* (curba de măsură a coeziunii agregării). Între indivizi, în cadrul unei agregări, există diferite distanțe. O măsură a acestei coeziuni este dată de relația:

$$C = (a - r) / d^2$$

( $d$  = distanță;  $a$  = forța de atracție iar  $r$  = forța de respingere -  $a=r$ ;  $C$  = coeziunea - în condițiile de mai sus coeziunea = 0; valoarea - pentru  $C$  arată o supraaglomerare iar indivizii se îndepărtează).

- *Coefficientul D e e v e y* (coeficient de aglomerare). Se referă la numărul de indivizi pe o anumită suprafață. Se poate aplica la specii sau formațiuni sedentare și care au conformație sferică (de ex. gale). Este dat de relația:

$$C_a = 2\pi r^2 \cdot N^2$$

( $r$  = raza individului sau a formațiunii;  $N$  = densitatea/unitatea de suprafață)

#### 6. Semnificația eșantionării, modelul de sampling și numărul optim de probe.

**A. Complexitatea eșantionării.** Un număr mare de specialiști în domeniu au abordat complex problematica eșantionării (sampling-ului) (aut.cit. în BINNS & NYROP, 1992), cu referire

la principiile de bază, avantaje, samplingul insectelor dăunătoare. Pe entomologii îi interesează de fapt două concepte practice: 1). **randomizarea** (aici trebuie să se ia cont de modelul de distribuție al populațiilor, deoarece în practică este mult mai important acest aspect decât niște probe luate absolut la întâmplare; de fapt cele mai reprezentative probe au fost obținute prin *eșantionare sistematică*); 2). **alegerea unității de probă și modul cum trebuie analizată, prelucrată și interpretată.**

În practica curentă entomologică sunt două direcții de bază pentru care se folosesc samplingul: a) - studii ecologice clasice la nivelul populațiilor și comunităților de insecte și b) - pentru elaborarea deciziilor în programele curente de monitoring și management ale insectelor dăunătoare. Deși punctul b. are importanță deosebită deoarece implică cunoaștere despre biologia, ecologia și comportamentul dăunătorului, tipul și caracteristicile daunelor, studiul densității populațiilor, stabilirea pragurilor economice, evaluarea dezvoltării în funcție de factorii de mediu, elaborarea modelelor de prognoză și control a populațiilor, programul deciziilor de intervenție, nu doresc totuși să dezvolt, acum și aici, aceste aspecte.

**B. Modelul de eșantionare (sampling).** Accentuez de la început că modelul de sampling este în funcție de obiectivul programului de cercetare. Dacă de ex. se doresc să se obțină *densitatea medie* pentru studiul ciclului de viață, atunci varianța (dispersia) se minimalizează. În schimb, dacă se studiază *modelul de distribuție*, o varianță mică nu este suficientă. Pentru a obține date cât mai revelatoare și un optim de informație este necesar să fie respectate câteva cerințe:

- Eșantioanele să fie luate randomizat (la întâmplare), cel mai corect fiind alcătuirea unui *plan de prelevare* (folosind chiar tabelele cu numere întâmplătoare sau o carte de telefon de unde se aleg ultimele două numere !), deoarece în câmp cercetătorul este tentat să aleagă sau să prefere anumite puncte și locuri de eșantionare. Nu se recomandă folosirea metodei de alegere la întâmplare a locului de eșantionare "absolut la întâmplare" (prin aruncarea unui obiect, din orice loc) fără a cunoaște comportamentul speciilor

- Se preferă *metoda randomizată stratificată* (pe orizontală și verticală) în acest caz, suprafața de probă fiind divizată în subunități de aceeași mărime. Această metodă asigură o acuratețe foarte mare în situația *estimării nivelului populațiilor*. Totuși, o estimare corectă și o atenuare a erorilor de eșantionare poate fi obținută prin *probe adiționale (suplimentare)*. Pentru a marca corect o agregare (*cartarea distribuției în aggregate*) se recomandă ca pe lângă proba randomizată, luat într-un anumit loc, să se ia încă o probă, la o distanță fixată în teren, de către cercetător. Dacă habitatul este stratificat iar specialistul grupului investigat, cunoaște comportamentul speciei, se pot elimina unele strate (nivele) unde s-ar găsi foarte puțini indivizi. De exemplu, la anumite specii de fluturi, distribuția pupelor în sol este în jurul bazei tulpinilor plantelor gazdă iar o estimare satisfăcătoare a nivelului s-a obținut dacă eșantionarea s-a limitat numai la această zonă (dacă operațiunea se face pentru mai multe grupe, atunci aceste nivale nu se elimină);

- Este necesar ca *programul de eșantionare să fie sistematic* (probele să se ia la intervale fixe în timp și spațiu). Logic, aici există cele 3 situații: un program redus (cu puține probe și luate nereprezentativ pentru stadiul de dezvoltare și evoluție al populației), un program mediu și un program complet sau foarte extins (probe luate săptămânal sau chiar zilnic, pe toată durata de dezvoltare a speciei în câmp). Prima variantă nu se recomandă dar nici ultima care, deși corectă, va afecta rentabilitatea economică, implicând cost și efort, ridicate. Important este obținerea unui *optim de informație*, la un cost moderat și efort mediu. De aici apare necesitatea studiilor preliminare, alegerea a cât mai multe puncte de probă din cele mai diverse habitate, binențeles, îmbinate toate cu o foarte bună cunoaștere a biologiei, ecologiei și comportamentului specific, dar și în funcție de scopul cercetărilor. Astfel în cercetări personale de ecologie la specii de lepidoptere noctuide, prin folosirea datelor de captură de la capcanele feromonale (aici considerate eșantioane), date cât mai relevante despre nivelul populațiilor, afinitate și diversitate s-au obținut prin folosirea datelor de captură de pe toată durata de zbor a speciei în natură, corelat și cu comportamentul larvar, statutul reproductiv (pe baza datelor de la capcanele luminoase) și relațiile cu factorii ecologici. În schimb, modelul de distribuție al populațiilor poate fi corect stabilit pe

un interval de 10-12 zile, la nivelul zborului 2 pe intervalul simetric maximului din curba de zbor. Si aici insă intervin alți factori (tehnici, comportamentul locomotor - dispersie, migrație) care pot denatura modelul populațional.

Se mai poate aplica programul randomizat sistematic repetabil, pentru un același habitat sau microhabitat, pe perioade variabile de timp:

- Sunt situații când eșantionarea se face prin luare în considerare a unei stări de fapt care nu jine de cercetător (ex. înregistrarea boabelor de fasole sau de porumb, atacate de gărgărije sau fluturi se face prin numărarea orificiilor; boabele de grâu care conțin larve de gărgărije sunt mai luminoase decât cele curate);

- Când se face eșantionare numai pentru un anumit stadiu al insectei este deci necesar să fie inclus vârful maxim al activității (dar din nou precizarea că totul este în corelație cu scopul studiului), dar este indicat a se corela acest aspect cu evoluția în descendență. Aici apare necesitatea corelării studiilor de autecologie în laborator și cu datele obținute în câmp, în cuști speciale;

- Perioada din zi când se face eșantionarea trebuie corelată cu ritmul circadian al activităților comportamentale (evident, fac excepție metodele speciale - capcanele, care sunt montate permanente) și cu deplasările pe verticală și orizontală.

**C. Experiențe de eșantionare. Tipuri și metode de eșantionaj.** În cele ce urmează prezint câteva detalii despre tipurile mai comune de eșantionaj și modul concret de operare în câteva cazuri.

a. **Sampling binomial (sistemul de marcare prezentă - absență).** Consta în a înregistra proporția de probe (eșantioane, subdiviziuni) în care se găsesc indivizi unei specii și nu numărul de indivizi/probă. Uncori, acest sistem se complică dacă se folosește combinat cu samplingul secvențial sau în situația folosirii a mai mult decât un singur plan de eșantionare. Totuși modalitatea aceasta s-a dovedit mai eficientă decât samplingul secvențial, nepractic pentru monitoring în livezi comerciale. Concret, luăm cauzul înregistrării proporției de frunze infestate cu acarieni, dintr-o lucrare reprezentativă (HEPWORTH & MACFARLANE, 1992b). Aici a fost dezvoltat un plan de sampling sistematic, cu mărime fixă a probei.

Metoda de lucru: La intervale de 9-10m s-a recoltat câte o frunză matură sau s-au luat câte 5 frunze de pe un rând de 40-50m, la interval tot de 4 rânduri. Frunzele (în jur de 180) s-au adus în laborator, s-au examinat, s-a notat numărul de indivizi, s-a estimat media ( $m$ ) și varianța ( $s^2$ ) și s-a calculat proporție de frunze lăță acarieni ( $p$ ). Datele, prin logaritmare, se reprezintă grafic, obținând o reprezentare ca în Fig. 38, cu marcarea ecuației distribuției, care este de tipul:  $\ln(m) = a + b \ln(-\ln p)$ . În lucrare, autorii specifică faptul că au făcut mai multe eșantionări și în alte suprafețe și analizează comparativ valorile parametrilor din ecuație. Pornind de aici, autorii estimatează densitatea populației din câmp (detalii în numărul următor referitor la dispersie, estimarea densității și nivelului absolut și relativ al populațiilor).

În cercetări personale făcute asupra estimării nivelului populațiilor de larve de *M. brassicae* și corelarea cu nivelul populației de adulți, acest model de eșantionare a fost unul dintre cele 6 analizate comparativ, obținând în final un model propriu<sup>1</sup>. Doar orientativ prezint aici, în mare, câteva aspecte al planurilor de eșantionare folosite pentru *M. brassicae* în vederea stabilirii acestor corelații. În studii preliminare (1984-1986) s-au ales suprafețe de 10 x 10m în vecinătatea fiecărei capcane, toate plantele de cultură (varza) fiind controlate (din 4 în 4 zile, din momentul apariției primului adult, până la impupare), notând nr. ponte, nr. ouă/pontă, nr. larve/vârste și apoi calculând frecvența de atac, frecvența plantelor neatacate, pentru fiecare observație estimând varianța și media. Începând din 1986 am mai introdus 5 metode (cea prezentată aici, lotul de 10

<sup>1</sup> STAN, GH., CORCIU, I., CRIȘAN, AL., CHIȘ, V., POP, L. M. *Mamestra brassicae* L. (Lepidoptera: Noctuidae): Relations among egg masses density, larval infestation, damage level and the number of male moths caught in light and sex attractant traps, in Cluj area (Transilvania, Romania) (in press).

$\times 10m$  și încă patru) analizate comparativ. Cele 4 au fost: a. - verificarea a unui nr de 300 plante, alese randomizat în jurul fiecărei capcane; b. - verificarea a către 10 plante, din 5 în 5 rânduri, din 20 în 20m pe rând, până la însumarea a 300 verze (aditional s-a folosit aceeași metodă dar până la însumarea a 100 verze; diferențele au fost nesemnificative); c. - verificarea a căte 200 plante (plus încă 100 plante) de pe 10 rânduri, cu 20 plante pe rând (respectiv, 10), prima fiind

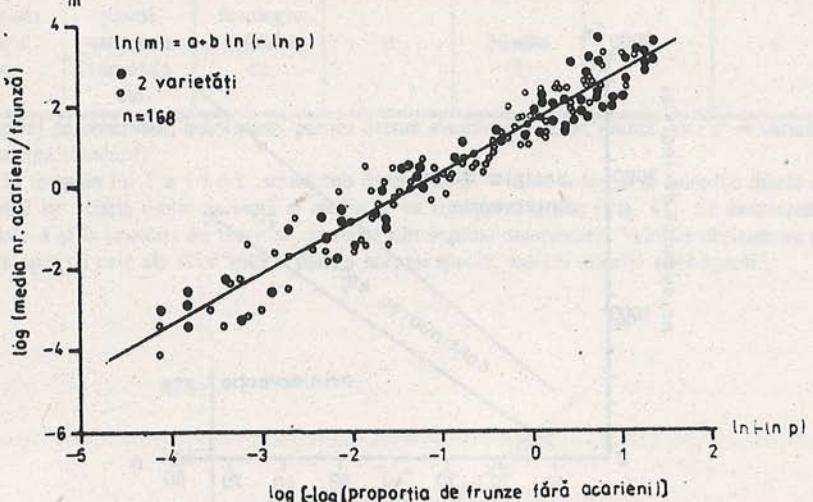


Fig. 39. Un model de reprezentare a rezultatelor unui sampling binomial, dând ca exemplu relația dintre media numărului de indivizi/frunză și proporția de frunze fără indivizi, pentru două varietăți de plantă gazdă. Dreapta de regresie este dată de relația  $\ln(m) = 1,62 + 1,24 \ln(-\ln p)$  ( $r^2 = 0,94$ ;  $n = 168$ ). Datele se referă aici la acarieni (după HEPWORTH & MACFARLANE 1992) dar modelul poate fi adaptat la orice specie de insectă, evident asociat cu tipul de comportament.

aleasă la întâmplare și apoi controlate toate, în ordine; d. - pe baza unui plan inițial făcut în laborator (pe baza nr. întâmplătoare), pentru fiecare control, s-au stabilit rândurile și poziția a 10m din fiecare rând unde s-au controlat toate plantele, operațiunea fiind efectuată între 2 și 20 repetiții (și poziția acestora s.a ales tot randomizat). Optimul de informație s-a analizat pentru 6 situații: 2, 4, 8, 12, 16 și 20 rânduri (pentru porțiunea de 10m, evident). La un număr de 25-30 plante/10m, în final am constatat că o verificare a 240 plante (în medie)(deci 8 rânduri  $\times$  10m/rând) este suficient și optim pentru a stabili că mai corect frecvența de atac și estimarea densității. Aceasta a fost aleasă apoi ca metodă pentru studiile din 1987-1989 pentru estimarea abundenței populațiilor locale ale acestui dăunător. Este evident că volumul de muncă a fost enorm dar necesar pentru cercetări eficiente. Probabil că în situația economiei de piață dacă s-ar lua în atenție și rentabilitatea, asemenea efort nu merită, mai ales în condițiile în care în final nici nu se poate utiliza în practică metodologia !. Așa cum am mai arătat, nu toate programele de eșantionare sunt la fel de complicate și depinde de scopul pentru care sunt făcute.

b. Sampling secvențial. Este o prelevare aleatoare de probe (unități statistice de eșantionaj) și cu o mărime a acestora, dependentă de observații făcute permanent în câmp. Unii autori au constatat, într-un anumit context (sampling pentru luarea deciziilor în monitoring și management la insecte dăunătoare), că metodele secvențiale sunt mai rentabile, costul este mai mic (uneori cu până 50%) comparativ cu metoda probelor de mărime fixă (BINNS & NYROP, 1992). Se constată că tot mai mult eficiența cercetărilor este legată de *costul obținerii informației și de ușurința cu care procedura poate fi dezvoltată și folosită pe scară largă*.

Testul secvențial a fost inițial definit după o ipoteză statistică, după care, la fiecare

observație, se pot lua 3 decizii: se acceptă ceea ce ne-am propus, se respinge sau se continuă eșantionajul (până ce cercetătorul va alege una din primele două; continuarea eșantionajului se poate face și prin observarea unei unități de eșantionaj, luată suplimentar). Funcția care permite această alegere implică numărul cumulat de indivizi numărați ( $T_n$ ) pe de pe unitățile de eșantionare, deci în funcție de  $n$ . Reprezentarea grafică (Fig. 40) va permite și definirea zonelor

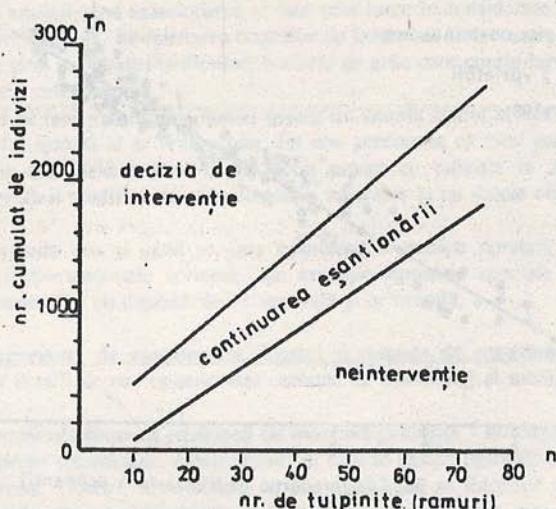


Fig. 40. Planul de eșantionare secvențială decizională, pentru specia *A. pisum*. Pragul de inervenție a fost fixat la 30 indivizi per plantă. Riscul:  $\alpha = 5\%$  (după BADENHAUSSER 1989).

care corespund cu deciziile ce trebuie luate. Se poate spune că este vorba de un **eșantionaj secvențial decizional** (deoarece, pornind de la stabilirea nivelului de densitate al populației în raport cu niște praguri teoretice se poate acționa asupra speciilor dăunătoare). Si în în acest caz am selectat un mod de lucru pentru afide, mai ușor de analizat dar bine documentat științific (după BODENHAUSSER 1989).

Metoda de lucru:

- **pentru un eșantionaj sistematic:** o parcelă se împarte în patrate, pe o întindere ușor de reperat în spațiu pentru a putea aprecia dispersia. Ex. O parcelă de  $1000m^2$  ( $50 \times 20m$ ) s-a împărțit în 50 de patrate de  $20m^2$  ( $5 \times 4m$ ). În fiecare bloc sau plot, pentru plantele de cultură (aici mazărea) s-au ales la întâmplare, 4 tulpini în porțiunea terminală (numărul este în funcție de atac, nivelul de aglomerare și mărimea aglomerării) și s-au făcut observații pe o durată de 3 săptămâni după înflorire (deci atenție, trebuie cunoscute biologia și comportamentul speciei pentru intervenție în momentul optim; se cunoaște că afidele au o repartiție spațială de tip aggregativ, fiind de fapt poate reprezentantul cel mai fidel al acestui model). Se face apoi numărarea (fiecare tijă se scutură pe un substrat mobil) iar apoi datele se analizează printr-o metodă neparametrică a tratamentului grilelor.

- **eșantionaj prin sondaj aleator:** are la bază ipoteza repartitionei la întâmplare a agregatelor de afide din câmp, sondajul fiind făcut în parcelele cercetate. În felul acesta se poate estima caracteristica spațială a agregării pentru fixarea procedurii secvențiale. În acest caz, timp de 7 ani s-au eșantionat mai multe varietăți de mazăre în diferite stadii fenologice. Mărimea eșantionării a fost variabilă (de la 25 la 75 tulpini alese la întâmplare).

Cu datele obținute se face un tabel (valabil și pentru alt tip de cercetări similare), cu următorul conținut:

Anul, perioada zborul etc	Profilul planta varietatea suprafața etc	Stadiul fenologic Mărime etc	n	Media ( $\bar{x}$ )	$s^2$	s
------------------------------------	--	---------------------------------------	---	------------------------	-------	---

(n = numărul de observații, eșantioane, pentru fiecare eveniment - zbor, plantă, etc;  $s^2$  = varianță; s = abaterea standard)

În formula lui T a ylor scrisă sub formă  $\log s^2 = a + b \log \bar{x}$  se introduc datele din acest tabel iar relația dintre varianță și medie se va reprezenta grafic (Fig. 41). Se estimează și variabilele a și b (metoda de lucru se cunoaște din studiile anterioare). Valorile obținute se pot compara apoi cu cele ale altor autori, pentru aceeași specie, sau cu valorile altor specii.

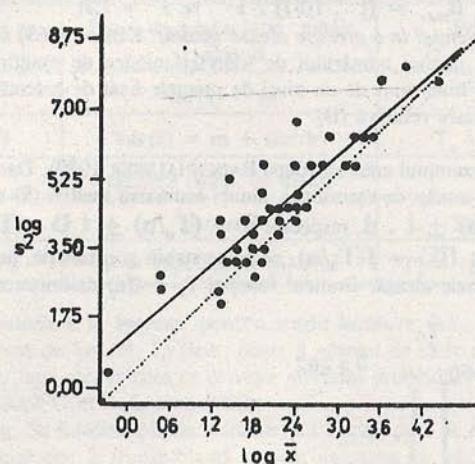


Fig. 41. Reprezentarea relației dintre varianță și medie (log) în cazul sondajului pentru specia *A. pisum*, în diferiți ani consecutivi de cercetare (după BADENHAUSER 1989).

Se alege un prag de intervenție (aici 30 indivizi/tulpiță, prag intenționat supraestimat; la *M. brassicae* a fost de ex. 2,5 larve/plantă).

Modelul de relație a fost:  $s^2 = 4,05 \bar{x}^{1,47}$ . Calculul curbelor de decizie ( $C_1$  și  $C_2$ ) s-a făcut estimând media la punctul  $\theta = 30$ , relația fiind pentru  $T_{\theta_0}/n \pm$  intervalul de confidență asociat la această estimare:

- pentru  $C_1$  -  $T_{\theta_0} = n\theta + t \sqrt{n} \cdot 4,05 \theta^{1,47}$
- pentru  $C_2$  -  $T_{\theta_0} = n\theta - t \sqrt{n} \cdot 4,05 \theta^{1,47}$

( $T_{\theta_0}$  = nr. cumulat de indivizi de pe n tulpițe observate și calculat pentru fiecare valoare a lui n, cu  $\theta=30$ ; t = testul "t" pentru  $\alpha/2$  și  $n-1$  grade de libertate)

Curbele de decizie sunt date în Fig. 40, pentru un risc de  $\alpha = 5\%$ . Pe măsură ce se eșantionează, punctele, corespunzător coordonatelor ( $T_{\theta_0}$  și  $n$ ), se vor trece pe grafic. Această procedură se face atunci când densitatea este ori prea scăzută, ori prea mare. Dacă la populația aceasta de afide, media  $< 7$  indivizi/tulpiță sau media  $> 50$  ind./tulp. se vor face observații

doar pe 10 tulpiji (în timp ce mărimea de eșantionare în procedură clasică este de 25 tulpiji). În cazul în care nu se poate lua rapid o decizie, mărimea eșantionului maxim se va calcula fixând avans la intervalul de confidență (I) asociat la estimarea mediei în punctul  $\theta$  (de ex. dacă intervalul de confidență este fixat la 6 indivizi,  $n_{max} = 67$  tulpiji, iar media este estimată  $T_{67} / 67 \pm 6$ .

**Completare.** IWAO (1975) este la originea acestei metode prin estimarea intervalului de confidență (I) asociat cu estimarea mediei ( $x$  sau  $m$ ) ca parte a unui eșantion de mărime  $n$ :

$$I = (t \cdot s) / \sqrt{n}$$

( $s$  = abaterea;  $t$  = student test la  $n-1$  gr.lib. -  $\alpha/2$ ). Pornindde aici, dacă  $s$  este o funcție cunoascută a mediei estimate, este posibil de stabili două relații corespunzătoare c curbelor care vor marca oprirea eșantionajului. Luarea deciziei se va face prin compararea densității observate, cu densitatea critică, plus-minus I, când mărimea de eșantionaj este  $n$ . Deci dacă  $T_n$  este numărul cumulat de indivizi când densitatea este  $\theta$  și numărul de unități de eșantionaj este  $n$ , iar varianța ( $s^2$ ) este în funcție de medie ( $\bar{x}$ ) ( $s^2 = f(x)$ ), atunci cele două curbe (prezentate și mai sus) sunt date de relațiile:  $C_1 - T_n = n \cdot \theta + t \sqrt{n f(\theta)}$  și  $C_2 - T_n = n \cdot \theta - t \sqrt{n f(\theta)}$ . Procedura de sampling poate fi oprită atunci când densitatea observată este apropiată de  $\theta$ , adică:

$$n_{max} = [t^2 \cdot f(\theta)] / I^2 \text{ iar } I^2 = I(\theta).$$

- *Eșantionajul secvențial la o precizie aleasă (fixată).* KUNO (1969) a propus o procedură secvențială de estimare a mediei numărului de indivizi/unitatea de eșantionaj iar criteriu de întrerupere al eșantionării fiind atins de un nivel de precizie fixat de cercetător în termenii unei erori absolute ( $d$ ) sau eroare relativă ( $D$ ).

Continuăm tot cu exemplul anterior (după BADENHAUSSER 1989). Dacă  $T_n =$  nr. cumulat de indivizi numărăți pe  $n$  unități de eșantionaj, atunci estimarea mediei ( $\bar{x}$ ) este:

$$\bar{x} = (T_n/n) \pm t \cdot d, \text{ respectiv } \bar{x} = (T_n/n) \pm t D \cdot (T_n/n).$$

Dacă există o relație  $s^2 = f(\bar{x}) = f(T_n/n)$ , se poate stabili și o funcție, pe baza cărei se poate opri intervenția, la o precizie aleasă. Graficul funcției  $T_n = f(n)$  delimităază de asemenea două zone (Fig. 42).

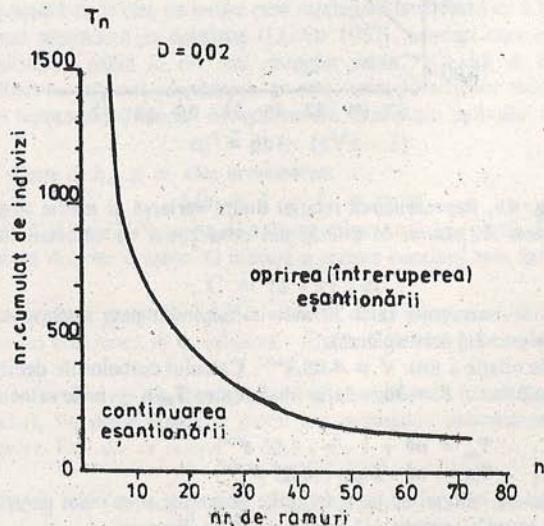


Fig. 42. Planul de estimare secvențială la precizia aleasă a mediei numărului de indivizi de *A. pisum*, per plantă. Precizia relativă,  $P_o = 0,20$ ;  $\alpha = 5\%$  (după BADENHAUSSER 1989).

In această figură, modelul de descriere al populației este ca și cel precedent iar funcția care va permite oprirea intervenției (tot pentru afidele de la mazăre), are următoarea formă:

$$T_n^{-0.53} = \frac{D^2}{4.05n^{-0.47}}$$

( $T_n$  = nr. cumulat de indivizi/ n tulipini observate; D = precizia relativă aleasă pentru estimarea mediei)

In figura de mai sus spațiul de dedesubt corespunde unei zone unde estimarea mediei prin  $T_n/n$  nu asigură precizia cercetătorului, iar opus, deasupra curbei, este cealaltă zonă. Fiecare punct al curbei este asociat exact cu precizia fixată. Dacă precizia absolută este fixată la d, atunci

$$d = \sqrt{f(T_n/n) / n},$$

iar dacă precizia relativă este fixată la D, atunci  $D = \sqrt{n \cdot f(T_n/n) / T_n^2}$ . Funcțiile  $T_n = f(nD)$  pentru principalele relații dintre varianță ( $s^2$ ) și medie ( $\bar{x}$ ), sunt prezentate în Tabelul 17.

**Tabelul 17**

Relațiile pentru calcularea curbei care marchează oprirea în planul de eșantionare secvențială (după BODENHAUSSER 1989).

Distribuția sau legea de probabilitate	Tipuri de relație între medie și varianță	Relația $T_n = f(n)$ D = fixat de cercetător
Poisson	$Var(\bar{x}) = m$	$T_n = 1 / D^2$
Binomial negativă	$Var(\bar{x}) = m + (m^2/k)$	$T_n = 1 / [D^2 - (1/nk)]$
Normală	varianță indep. de medie	
Legea lui Taylor	$s^2 = a\bar{x}^b$	$T_n^{b-1} = D^2 / an^{1-b}$
Relația lui Iwao	$s^2 = (a_2 + 1)\bar{x} + (b_2 - 1)\bar{x}^2$	$T_n = (a_2 + 1) / [D^2 + (1 - b_2)/n]$

**Model de eșantionare la insecte, pentru stadii imature (ou, larvă, pupă).** Poate fi folosit pentru orice specie de insectă. Evident, poate fi adaptat de către cercetător, în funcție de comportamentul speciei, mai ales în ceea ce privește mărimea probelor. Prezentăm aici un model pentru stadiul de ou, (după GOH & LANGE 1980).

Metoda de lucru. Se folosesc plante netratate. La fiecare dată de eșantionare s-au selectat câte 20 plante. S-au luat cite 2 frunze/plantă și s-au examinat în laborator (când punta este suficient de mare, numărarea se poate face direct în câmp). Adițional de pe fiecare plantă s-au mai selectat 2 frunze. Pentru determinarea distribuției de ouă/plante, de pe 5 ramuri laterale se adună frunzele în 4 repetiții (în funcție de mărimea plantei - mică, mijlocie, mare - de pe căte 3 plante s-au luat 5, 8 și respectiv 14 frunze). Similar s-a procedat și pentru larve. Pentru pupe, s-a eșantionat litiera de sub plante. Plantele din câmp se mai pot împărți și în categorii de vîrstă (pentru specii cu durată de viață de mai mulți ani). Se alege o anumită suprafață cu un număr corespunzător de plante/suprafață (ex. 20), pentru fiecare clasă de vîrstă. În acest caz, varianța ( $s^2$ ) se va calcula prin combinarea datelor de la toate clasele (deci 60 plante, dacă au fost 3 clase de vîrstă; fenomenul este similar dacă se aleg soiuri sau alți parametri). Frecvența distribuțiilor pentru indivizii numărați cu fiecare ocazie de eșantionaj au fost analizate pentru a determina aproximativ modelul de distribuție (folosind relația Taylor sau testul  $\chi^2$  - vezi Bul.inf.Soc.lepid.rom. 4(2): 109 /1993). Pentru distribuția binomial negativă, mărimea optimă a probei a fost calculată după ecuația din tabelul după KARANDINOS (1976), prezentat aici în cap. II. Pentru distribuția binomial negativă, valoarea lui k a fost calculată după relația  $N \ln(1 + \bar{x}/k) = \Sigma (A_x/k + \bar{x})$  (SOUTHWOOD 1966). În final, datele se pot aranja într-un tabel cu următoarele rubrici:

Data obs.	Câmpul (clasa, etc)			larve (ouă)/pl.*		$s^2$	$s^2/\bar{x}$	$k^{**}$	$n$
	I	II	III	$x \pm ES$	Interval				

\* bazat pe un număr de .... plante cu ....frunze/plantă;

\*\* parametrul k al distribuției binomial negative ( $P > 0,05$ ):

#### D. Numărul de eșantionări per unitate de habitat.

##### a. Subdivizarea habitatului.

Dacă distribuția populației în habitat este predispusă spre anumite subdiviziuni dar probele sunt luate la întâmplare (randomizat) vor apărea erori sistematice. Acestea pot fi înălțaturate printr-un sampling cu un număr de probe caracteristice pentru fiecare subdiviziune, sau se ia fiecare parte separat (unități de probă), se eșantionează și la urmă se analizează rezultatele. Subdiviziunile acestea evidențiază variație în funcție de habitat, de comportamentul speciei și de scopul acțiunii. Trebuie înținut astfel cont de marea variabilitate existentă în natură. Dacă am aminti aici doar un aspect referitor la eșantionarea pantei, date de specialitate au înregistrat variații ale numărului, în funcție de copac, înălțimea și consistența coroanei, nivelul din coroană, soi, poziția frunzelor, la specii ca *Archips argylospillus* sau *Choristoneura fumiferana*. Fenomenul este valabil și pentru specii care depun pontă pe diferite legume (ex. *Mamestra brassicae*). În toate aceste situații este necesară eșantionarea la mai multe nivele, orizontal și vertical, prin alegerea unui număr optim corespunzător de unități de probă, randomizat și cu păstrarea unor principii general recunoscute (pe căt posibil planul de randomizare să fie stabilit în laborator, probele să fie mai multe și mici decât puține și mari, să se facă eșantionări suplimentare în timp și spațiu). Deci, în programul de eșantionare se alege habitatul sau habitatele preferențiale, fiecare din acestea se va împărtăși în subunități (fiecare subdiviziune având mărimea corelată cu comportamentul speciei studiate (pentru afide habitatul este planta gazdă iar subunitățile sunt tulipinile, frunzele; pentru larvele de *M. brassicae* habitatul este cultura de varză, iar subdiviziunea este varza)). Vom vedea în continuare că în funcție de tipul de sampling (binomial, secvențial, ocazional), pentru un optim de informație există un raport optim între numărul de habitate și subdiviziuni analizate.

##### b. Numărul de probe pe o subdiviziune.

De exemplu, pentru a determina numărul de probe/plantă ( $n_s$ ), varianța pentru probele luate de pe aceeași plantă ( $s^2_s$ ), acestea trebuie comparate cu dispersia (varianța) probelor luate la diferite plante ( $s^2_p$ ), implicând costul de sampling pe o aceeași plantă ( $C_s$ ) și efortul eșantionării în mișcarea de la o plantă la alta ( $C_p$ ):

$$n_s = \sqrt{\frac{s_s^2}{s_p^2} \cdot \frac{C_p}{C_s}}$$

Dacă varianța între plante ( $s^2_p$ ) este o sursă majoră de variație și în afară de cazul în care costul de eșantionare este mare,  $n$  va fi de ordinul lui 1 sau  $< 1$  (în practică = 1). De exemplu, pentru diferite specii, cum sunt: *Choristoneura fumiferana*, *Operophtera brumata*, *Plutella xylostella*, *Pieris rapae* sau *Gilpinia hercyniae*, *Recurvaria starki*, *Brevicoryne brassicae*,  $s^2_p > s^2_s$ . La multe din cazuri,  $s^2_s$  a fost mică și s-a luat numai o probă/plantă sau per un anumit nivel al plantei (frunză, ramură). Pentru speciile dătoare la măr,  $s^2_s$  a fost mai mare iar în acest caz s-au luat căte 7 probe. Adesea, cu o considerabilă economie în cost, efort redus și fără pierderi în acuratețea evaluării populației (dar totuși cu pierderi la nivelul erorii de eșantionare), s-a obținut prin luarea randomizată a unui număr de subprobe (subeșantioane) care au fost cântărite, s-a înregistrat mărimea și s-a măsurat înainte de sortare și numărare (acest lucru este necesar acolo unde procesul de eșantionare este complex).

##### c. Numărul optim de probe dintr-un habitat.

Numărul de probe depinde de gradul de precizie cerut și tipul de studiu efectuat. În cadrul unui habitat ecologic omogen, numărul de probe cerut ( $n$ ) va fi:

$$n = (t \cdot s) / D \cdot \bar{x}$$

( $s$  = eroarea standard;  $t$  = cantitate dependentă de numărul de probe și se obține din tabele - pentru mai mult de 10 probe,  $t = 2$  la un nivel de 5%;  $D$  = nivelul de acuratețe cerut și exprimat decimal, de la 0 la 1).

Să ne reamintim cele două relații prezentate în Bul. inf. 4(2)/1993:

$$n = s^2 / (D^2 \cdot \bar{x}^2)$$

( $D$  = asigurarea aleasă - de ex. pentru 95%,  $D = 0,05$ , dar  $D$  poate fi și calculat pentru a avea informații optime:  $D = \sqrt{(s^2 / n) / \bar{x}}$ ).

Când samplingul este nevoie să fie făcut la două nivele (de ex. eşantionarea pontelor sau larvelor la nivel de frunze și apoi la nivel de copac), atunci numărul de unități de pecare ar fi necesar să se ia probe este dat de relația:

$$N_t = \frac{\frac{s^2}{n} + s^2_p}{\frac{n_s}{(\bar{x} + D)^2}}$$

( $n_t$  = nr. probe în cadrul aceleiași unități de habitat, calculat ca la punctul anterior;  $s^2_n$  = varianță în cadrul aceleiași unități de habitat;  $s^2_p$  = varianță între unitățile de habitat sau altfel spus, dispersia între plante;  $x$  = media per probă;  $D$  = aceeași semnificație ca sus).

Dacă dispersia populației este bine descrisă de o distribuție binomial negativă, adică în agregate ( $s^2 > x$ ), numărul de probe este dat de relația:

$$N = \frac{\frac{1}{\bar{x}} + \frac{1}{k}}{D^2}$$

$k$  = parametrul de dispersie al distribuției binomial negative (în agregate), despre care am amintit anterior. Cum se calculează  $k$ ? Există mai multe metode, dar aici prezint doar 3:

a).  $k = \bar{x}^2 / (s^2 - \bar{x})$  sau se calculează din următoarele două ecuații:

b).  $\log(N/n_0) = k \cdot \log[(1 + \bar{x}) / k]$

c).  $N \ln [1 + (\bar{x} / k)] = \sum [A_i / (k + i)]$

( $\ln = \log$ ;  $A_i$  = suma frecvențelor unităților de probă conținând mai mult de  $i$  indivizi; ex.  $A_6 = \sum (f_7 + f_8 + f_9)$ ).

Un alt tip de program de sampling constă în măsurarea frecvențelor unei specii sau eveniment (ex. frecvență de apariție a urmelor de atac larvar pe frunze; frecvență de apariție a galelor, etc). Înainte de estimare se vor lua toate probele necesare obținând astfel o valoare aproximativă a probabilității (ex. dacă într-un asemenea control preliminar s-a găsit că 25% din frunze sunt atacate, probabilitatea = 0,25). Numărul de probe va fi dat de relația:

$$N = (t^2 \cdot p \cdot q) / D^2$$

( $p$  = probabilitatea;  $q = 1 - p$ ;  $t$  = entitate dependentă de numărul de probe, fiind obținută din tabele;  $N > 10$  probe,  $t = 2$  la un nivel de 5%;  $D$  = nivelul de acuratețe cerut, poate fi exprimat decimal, de la 1 la 10)

Datele de specialitate (ex. KARANDINOS 1976) indică diferite formule (Tabelul 18) referitoare la numărul optim de probe, corelat cu tipul de distribuție și în funcție de 3 modalități de estimare a încrederii în date (deaiii ajutătoare în Bul. inf. Soc. lepid. rom., 4(2): 97-99, 1993, referitoare la *numărul de probe și indicele de confidență*).

Estimarea parametrilor unei populații din câmp este o parte indispensabilă a multor proiecte ecologice și de management. Un lucru rămâne totuși logic: cu cât este mai mare numărul de probe, cu atât estimarea nivelului corect al unei populații, este mai bună. Dar, aşa cum am mai arătat, "costul" fiind prea ridicat se renunță și la număr foarte mare de probe și la un studiu pe suprafețe foarte mari. În acest caz trebuie găsit un **număr optim de probe** și care să asigure totuși **încredere** în datele noastre. În tabelul de mai sus tocmai acest lucru este prezentat: formule care să definească această încredere ( $m$  = media;  $s$  = abaterea standard;  $s^2$  = varianță;  $k$  =

parametrul de dispersie:  $C$  = coeficientul de variabilitate, decis arbitrar pentru toate probele;  $D$  = proporția fixată pentru mediea  $m$ ;  $p$  = probabilitatea de succes, adică  $p = X/n$ . Dacă se ia o probă randomizată de mărime  $n$ , pentru o distribuție cu media  $m$  și varianța  $s^2$  atunci conform *Teoremei limitei centrale*, pentru un număr suficient de mare de probe (în jur de 30 sau mai multe), există următoarea relație probabilistică:

$$P[X - Z_{\alpha/2} \cdot (s/\sqrt{n}) < m < X + Z_{\alpha/2} \cdot (s/\sqrt{n})] = (1 - \alpha)$$

( $X$  = media probei;  $1 - \alpha$  este coeficientul de confidență;  $Z_{\alpha/2}$  = punctul de deasupra lui  $\alpha/2$  al unei distribuții standard normale; pentru  $1 - \alpha = 0.95$ , rezultă  $Z_{\alpha/2} = 1.96$ ).

În ceea ce privește distribuția binomială negativă, formula dată în tabel a fost obținută de autor prin substituiri succesiv. Anterior am prezentat și mai reamintim și aici, faptul că asemenea formule au fost date și de alți autori.

De exemplu SOUTHWOOD (1971) dă următoarea relație:

$$n = (1/m + 1/k) / D^2$$

(semnificația termenilor este aceeași dar este omis  $Z_{\alpha/2}$ ). O altă relație

**Tabelul 18**  
Formule pentru estimarea numărului optim de probe (după KARANDINOS 1976)

Tipul de distribuție al populației	Increderea în date este definită de:		
	Coeficientul de variabilitate ( $C_v$ )	Formule statistice. Intervalul de confidență este egal cu:	
		% al parametrului	un număr pozitiv fixat
Generală	$n = (\frac{s}{mC})^2$	$n = (\frac{Z_{\alpha/2}}{D})^2$	$n = (\frac{Z_{\alpha/2}}{h})^2$
Binomial negativă (agregate)	$n = \frac{\frac{1}{m} + \frac{1}{k}}{C^2}$	$n = (\frac{Z_{\alpha/2}}{D})^2$	$n = (\frac{Z_{\alpha/2}}{h})^2$
Poisson (întâmplătoare)	$n = \frac{1}{m \cdot C^2}$	$n = (\frac{Z_{\alpha/2}}{D})^2$	$n = (\frac{Z_{\alpha/2}}{h})^2$
Binomială	$n = \frac{q}{p \cdot C^2}$	$n = (\frac{Z_{\alpha/2}}{D})^2$	$n = (\frac{Z_{\alpha/2}}{h})^2$

este cea dată de PIETERS & STERLING (1974):

$$n = (1/m + 1/k) / D$$

KARANDINOS (1976), compară formula personală dată în tabelul de mai sus, cu aceste două

formule (în ordinea prezentată aici). Pentru valorile cunoscute ale parametrilor ( $m=3$ ;  $k=2$ ;  $D=0,1$ ;  $\alpha=0,05$ ;  $Z_{\alpha/2}=1,96$ ), valoarea numărului optim de probe a fost: 320, 83 și respectiv 8,3. Răspunsul corect este 320, cel obținut după relația din tabel. Evident, apare întrebarea, care formulă este corectă și trebuie folosită de către noi? O parte din răspuns poate fi dedus din ce am afirmat anterior: cercetătorul să aibă în vedere mereu, care este scopul cercetărilor. Ulterior, prin experiență câștigată, prin informație științifică în domeniu, va putea pune la punct un model de eșantionaj caracteristic comportamentului speciilor, repetabil în timp și spațiu, pe baza căruia se poate obține un optim de informație.

Gh. STAN  
Institutul de Cercetări Biologice  
Col. Entomologie experimentală  
Str. Republicii, nr. 48  
3400 CLUJ-NAPOCA

Apărut/Printed: 9.09.1996