

der har overvintret 2 eller flere gange. Sommergenerationen består af overlevende fra forårsgenerationen samt et stort antal nyklækkede individer. Levealderen blev omtalt. Af 109 mærkede biller, der blev fundet et til flere år efter første mærkning, havde 72 overvintret een gang, 19 to gange, 14 tre gange og 4 fire gange. Ovarieundersøgelser viste, at ovarierne efter at være brugt på ny udviklede sig, og *H. smaragdinus* havde således mindst to forplantningsperioder.

#### Populationsberegninger.

Overlevelseshæftigheden blev fundet på grundlag af ligningen  $N_t = N_0 \cdot e^{-x \cdot t}$ . Når maximum  $N_0$  sættes på et tidspunkt, hvor klækning er ophørt, og man således ved, at der ikke er nogen tilgang på grund af klækning, og der heller ikke kan ske nogen ind- eller udvandring som på st. 2, så må den afgang, der kan iagttages, skyldes dødeligheden, og der kan derfor fremstilles en overlevelseshæftighedskurve, der naturligvis kan udnyttes til at bestemme, hvor stort et antal mærkede dyr der er tilbage på lokaliteten, og derefter kan ligningen  $x = \frac{a \cdot (n+1)}{(r+1)}$  benyttes til bestemmelse af populationens størrelse.

Der kunne påvises en tilbagegang i bestanden fra år til år, som sandsynligvis skyldtes ændringen i vegetationen. En stærk tilbagegang i 1959 og 1960 skyldtes det varme og tørre år 1959. Sommergenerationen i 1959 udeblev næsten fuldstændig, hvilket naturligvis fik betydning for bestandens størrelse i 1960.

**Walter Hackman** nämnde att han använt Fischer-Ford's beräkningsmetod på basis av märkning vid uppskattning av populationsstorlek hos en lycosid-spindel och ansåg denna metod giva rätt tillförlitliga resultat med beaktande av dödligheten inom populationen.

#### F. Søgaard Andersen: *Statistisk behandling af antallet af dyr i prøver.*

Hvis man tæller børsterne på et bestemt organ på et stort antal insekter af samme art, finder man sædvanligvis, at de danner noget, der ligner en normal fordeling, d. v. s. hvis man sætter antallet af børster som abscisse og antallet af individer med det pågældende antal børster som ordinat, får man den bekendte, klokkeformede kurve. Det samme gælder, hvis man

måler eller vejer det samme organ, i hvert fald hvis man tager logaritmen til størrelsen eller vægten. Det er den normale fordeling, der er forudsætningen for de almindelige statistiske beregninger: spredning, middelfejl på middeltal og forskelle mellem middeltal, t-test, variansanalyse, regressionsanalyse o. s. v.

Det er altsammen tilnærmelser, men det fungerer godt, først og fremmest fordi variansen og spredningen er praktisk talt uafhængige af middeltallet.

Desværre er problemerne meget mere komplicerede, når det gælder antallet af dyr i prøver.

Må jeg minde om, hvor almindeligt det er, at man tager et antal prøver af samme størrelse og tæller dyrene for at bestemme tætheden af en eller flere dyrearter: bundprøver i havet og i søer, jordprøver, plankton-prøver, kvantitative ketsjninger, optælling af skadedyr på enkelte blade eller grene, fangst i et bestemt tidsrum med fælder, f. ex. lysfælder, suge-fælder, klæbefælder o. s. v.

I alle disse tilfælde får man fordelinger af hyppigheder af dyr, og disse fordelinger er altid skæve, og variansen tiltager altid med middeltallet. Altså, hvis man sætter antallet af dyr pr. prøve som abscisse og antallet af prøver med det pågældende antal dyr som ordinat, får man altid en usymmetrisk kurve med den lange hale til højre, i retning af stigende antal dyr pr. prøve.

Som eksempel kan jeg nævne, at jeg tog 12 prøver byg på et kornloft for at skønne, hvor mange kornsnudebiller der var. Hver prøve var på  $\frac{1}{2}$  kg, og antallet af kornsnudebiller fordelte sig sådan, at der var 8 prøver med 0, 2 med 1, 1 med 4 og 1 med 25 biller. 12 andre prøver fra samme loft indeholdt ikke en eneste bille. Andre prøveserier lå derimellem.

Hvad skal man nu stille op med sådanne resultater? Hvad kan man vide om antallet af kornsnudebiller på hele loftet, og hvordan skal man forestille sig billerne fordelt på loftet, for at det kan føre til sådanne resultater.

Lad os først se på, hvordan dyrene ville være fordelt, hvis det var helt tilfældigt.

Hvis sandsynligheden for, at et dyr er i en bestemt prøve, er  $p$ , og sandsynligheden for, at det er i området udenfor prøven er  $q = 1 - p$ , vil fordelingen af  $k$  dyr være i overensstemmelse med binomialfordelingen  $(p + q)^k$

med middeltallet:  $m = kp$

og variansen  $\sigma^2 = kpq$

Hvis man indsætter  $q = 1 - p$  og  $kp = m$  får man

$$\sigma^2 = kp(1 - p) = kp - kp^2 = m - m^2/k.$$

Hvis nu  $k$  bliver meget stor, og det er den, når hele området er stort, så er variansen lig med middeltallet, og så har man Poisson-fordelingen. Af og til finder man faktisk tilflyvende skadedyr, der fordeler sig på denne måde i prøverne.

Men i de allerfleste tilfælde er antallet af dyr i prøverne fordelt tilnærmelsesvis efter den negative binomial-fordeling:

$$(q - p)^{-k}$$

hvor  $q = p + 1$

$$m = kp \text{ og}$$

$$\sigma^2 = kqp = kp + kp^2 = m + m^2/k.$$

Den sidste ligning viser, at variansen stiger med kvadratet på middeltallet; spredningen er altså nærmest retliniet forbundet med middeltallet.

Endvidere viser ligningen, at jo mindre  $k$  er, des stærkere stiger variansen med middeltallet, og det hænger igen sammen med, at fordelingen er skæv;  $k$  er altså et mål for skævheden.

Det er klart, at på sådanne fordelinger kan man ikke anvende de almindelige operationer, som gælder for den normale fordeling, altså variansanalyse, regressionsanalyse, t-test o. s. v., fordi variansen er afhængig af middeltallet, og fordi ydre faktorerers indvirkning på de målte størrelser ikke er additive.

Og nu er jeg omsider nået frem til emnet for mit foredrag, nemlig: hvad skal man da gøre med sådanne tal?

Svaret er, at man kan gå to veje.

Den almindelige vej, som iøvrigt er den eneste, der er rigtigt gennemarbejdet, er at transformere tallene således, at de bliver tilnærmelsesvis normalt fordelt. Det medfører, at variansen bliver praktisk talt uafhængig af middeltallet og virkningen af ydre faktorer i hovedsagen bliver additiv, og så regner man på det, som om det var normalt fordelt.

Den bedste transformation er i de fleste tilfælde

$$x' = k^{\frac{1}{2}} \sinh^{-1} (x/k)^{\frac{1}{2}}$$

eller for små værdier af  $x$  (viz.:  $\bar{x} < 10$ )

$$x' = k^{\frac{1}{2}} \sinh^{-1} ((x + 0.5)/k)^{\frac{1}{2}}.$$

Man har en anden transformation, som er lidt simplere og næsten lige så god, nemlig

$$x' = \log (x + k/2).$$

Begge disse transformationer forudsætter, at man estimerer  $k$ , og det er ret tidsrøvende; der er derfor nogle, som bruger

$$x' = \log(x + 1)$$

og det går meget godt i mange tilfælde, hvis det kun drejer sig om en variansanalyse el. lign., men hvis middeltallet er lille og  $k$  er lille, går det ikke, og man får jo ingen oplysning om, hvor klumpet fordelingen er.

Hvis man beregner  $k$ , har man i hvert fald et mål for klumpningen, og har man taget prøveserier under forskellige forhold, f. ex. på forskellige lokaliteter eller til forskellige tider, kan man undersøge, om klumpningen er forskellig.

Det kunne nu se ud, som om problemet er løst: man skal bare tælle dyrene i serier af tilfældigt fordelte prøver, undersøge om serierne har fælles  $k$  (hvad de som regel har), og hvis det er tilfældet, da bestemme den fælles  $k$ , transformere observationerne efter formelen  $x' = k^{\frac{1}{2}} \sinh^{-1}(x/k)^{\frac{1}{2}}$  og derefter regne variationsanalyse eller regressionsanalyse, som om de transformerede observationer er normalt fordelt.

Men i mange tilfælde er det meget utilfredsstillende.

For det første er  $k$  ofte meget lille, og det medfører, at variansen både på observationerne og på  $k$  er meget stor, så at det bliver umuligt at konstatere forskelle på noget som helst.

For det andet: når man har fået at vide, at dyrene fordeles sig i klumper, vil man gerne vide, hvor store klumperne er, og hvor tæt de ligger, og man vil også gerne vide, hvorfor de klumper sig sammen.

Hvis man vil have mere at vide, må man forandre indsamlingsplanen.

Man kan simpelt hen tage større prøver. Det vil i hvert fald bevirke, at middeltallet stiger, og i næsten alle tilfælde vil  $k$  også stige med det resultat, at variansen på middeltallet falder, og med et stort middeltal og et stort  $k$  bliver transformationen mere effektiv.

Men de store prøver giver ikke flere oplysninger om dyrenes fordeling end de små, snarere tværtimod.

En af de veje, man kan gå, er at tage mange små naboprøver, som kan sammenlægges til en stor prøve. Hvis sådanne naboprøveserier fordeles tilfældigt ud over et område, får man en hel del at vide.

Boy Overgaard\*) har praktiseret det princip ved indsamling af larver af mitter (*Culicoides impunctatus*) i Lyngby Åmose i foråret. Han tog kvadratiske prøver af Sphagnum-bunden på  $10 \times 10$  cm<sup>2</sup>, men på den måde, at han tog  $4 \times 4$  sammenstødende prøver på 12 tilfældigt valgte steder, og i tilgift tog han 2 serier på  $8 \times 8$  prøver.

Hvis man ser bort fra den ene af de store serier, hvor der er mange larver i alle prøverne, er det en nogenlunde jævn blanding af små og store tal, fra 0 til 29. Tallene fordeler sig som en negativ binomial fordeling. Det viser, at larverne er fordelt i klumper. Og i kraft af indsamlingsmåden kan man nu vise, at klumperne er mindre end prøvestørrelsen. Det kan vises på følgende måde:

Man danner en prøveserie ved at udtage en tilfældig prøve af hvert af kvadraterne, og det gentager man, til man har 16 prøveserier, og så beregner man den fælles  $k$  for de 16 serier.

Derefter betragter man de 16 prøver indenfor hvert kvadrat som en prøveserie og beregner en fælles  $k$  for alle kvadraterne.

Hvis klumperne var større end prøve-størrelsen, ville tallene indenfor et kvadrat være mere afhængige end de tilfældigt udtagne fra hvert kvadrat, og så skulle den sidste beregning give større  $k$  end den første, men det er ikke tilfældet; de er nogenlunde lige store.

Derimod bliver  $k$  meget større for den ene af de store serier på  $8 \times 8$  prøver. Her er altså tale om en klumpning af anden orden, en sammenklumpning af klumperne.

En hypotese til forklaring af de to størrelser af klumper er nærliggende: De små klumper er ægportioner, og der kan naturligvis være lagt en eller flere portioner indenfor samme prøves område. De store klumper kan skyldes, at hunnerne søger sammen under æglægningen. Det er iagttaget for en anden art. Den store klump skulle så stamme fra en lun nat i en tid med maksimum af drægtige hunner, som er søgt sammen over et begrænset område, hvorimod æggene i det mere tyndt befolkede område er lagt enten på nætter med dårligt vejr eller på tidspunkter før og efter hunnernes maximum.

Vi har således fået nogle oplysninger, som den gammeldags indsamlingsmetode ikke kunne give. Men materialet er ikke fuldt udnyttet endnu. Thompson (1955) har udarbejdet en metode til

\*) se Statens Skadedyrlaboratoriums årsberetning 1961 og 1962.

en variansanalyse, hvor man undersøger variationen mellem prøver indenfor par, mellem par indenfor tetrader, mellem tetrader indenfor oktader o. s. v. Den analyse har bl. a. betydning til efterprøvelse af stokastiske modeller, som man kan opstille, hvis man har nøjagtige, biologiske oplysninger om arten.

Som eksempel på, hvordan sådanne biologiske oplysninger skal være indsamlet, kan vi tage mitterne (*Culicoides*) i Åmosen; for at bygge en stokastisk model for larvernes fordeling på fladen må man have følgende oplysninger:

1) Fordelingen af antal æg pr. portion og 2) af antal portioner pr. hun, samt 3) af den afstand hunnerne bevæger sig mellem hver æglægning.

Endvidere 4) fordelingen af antallet af drægtige hunner i tid og 5) af størrelsen af sværmene i forhold til antallet af drægtige hunner en given nat. Og endelig må man have 6) fordelingen af dødeligheden blandt æg og larver indtil observationstidspunktet, og 7) af den distance larverne bevæger sig bort fra æggene.

Hvis det lykkes at opbygge en stokastisk model, bliver man i stand til at angive sandsynligheden for forskellige observationer nøjagtigt, ikke med mere eller mindre grov tilnærmelse, som det nu sker ved hjælp af den normale fordeling.

Bartlett, M. S. (1960): Stochastic population models. (London). — Cassie, R. M. (1962): Frequency distribution models in the ecology of plankton and other organisms. — *J. Anim. Ecol.*, 31: 65—93. — Thompson, H. R. (1955): Spatial point processes, with application to ecology. — *Biometrika* 42: 103—115.

**Svante Ekholm:** Vilken litteratur föreslår föredragshållaren för våra biologer?

**Søgaard Andersen:** Regneskemaer for elementære statistiske analyser findes i A. Hald: Statistical theory with engineering applications (N. Y. & London 1952). Hvis det specielt drejer sig om den negative binomial-fordeling, anbefales de gennemregnede eksempler i C. J. Bliss & A. R. S. Owen: Negative binomial distribution with a common  $k$ . — *Biometrika* (1958) 45: 37—58.

**Lars Hedström:** *Sexuellt beteende inom familjen Dolichopodidae (Dipt.).*

Studierna över styltflugornas frieri- och kopulationsbeteenden har företrädesvis utförts i fält i flugornas naturliga miljö. De viktigaste tekniska hjälpmedel som komit till användning är en 16 mm smalfilmkamera Paillard-Bolex H16 Reflex och en