

Matematyka dla biologów — Zajęcia nr 9.

Dariusz Wrzosek

25 listopada 2024

Modelowanie rzeczywistości

Słowo ”**model**” ma kluczowe znaczenie w nauce, ale używane jest w różnych kontekstach. Modele konstruuje się po to, aby zrozumieć procesy przyrodnicze i przewidywać ich przebieg (np. prognozowanie pogody, rozkład wiekowy populacji, tempo wzrostu guza nowotworowego i.t.d.). Spotyka się trzy pojęcia, których granice są nieostre :

- **model statystyczny**—np. prosta (ogólnie krzywa) regresji jest modelem stat. określającym zależność między dwoma cechami, niekoniecznie związanymi przyczynowo-skutkowo, podobny charakter ma linia trendu liniowego uzyskana metodą najmniejszych kwadratów,
- **model symulacyjny (albo komputerowy)**—np. model agentowy tzn. algorytm, który tworzy w pamięci komputera wirtualną populację komórek według przyjętych zasad rozwoju,
- **model matematyczny**—np. równanie rekurencyjne określające stan populacji w k -tym sezonie w zależności od stanu w sezonie $k - 1$, albo model określający prawdopodobieństwo wystąpienia mutacji w danym fragmencie DNA w ustalonym przedziale czasu.

Modele matematyczne

Upływ czasu w modelu można reprezentować:

- **dyskretnie** za pomocą zbioru liczb naturalnych odpowiadających wybranym chwilom czasu;
- w sposób **ciągły** za pomocą zbioru liczb rzeczywistych.

Oba typy modeli spotykamy w ekologii, genetyce populacyjnej, kinetyce reakcji chemicznych, fizjologii.

Modele mogą być

- **deterministyczne** — określają (średni) stan jakiegoś układu w czasie,
- **losowe** (stochastyczne) — określają rozkłady prawdopodobieństwa, że dany układ znajdzie się w danym stanie.

Dobrze sprecyzowany model stanowi przedmiot badań matematycznych.

Dowodzi się twierdzenia charakteryzujące własności modelu.

W szczególności trzeba stwierdzić, czy nie jest on wewnętrznie sprzeczny (np. czy rozwiązania użytego równania w ogóle istnieją itp.).

Weryfikacja modelu

Ważnym zagadnieniem jest **weryfikacja modelu**, czyli porównanie tego co on przewiduje z rzeczywistością.

Jest to o wiele trudniejsze w przypadku modeli biologicznych niż np. w fizyce, ze względu na nieporównywalnie większy stopień złożoności układów biologicznych i ich naturalną zmienność.

Z tego powodu modelując procesy biologiczne nie należy liczyć na to, że można przewidzieć przebieg zdarzeń z dokładnością porównywalną z tą, którą osiąga się w fizyce lub technice.

W przypadku modeli biologicznych można być zadowolonym nawet gdy model matematyczny przewiduje tylko część cech układu rzeczywistego, np. pewne cechy jakościowe typu wzrost, spadek, okresowość, ale jest niedokładny pod względem ilościowym.

- Nie ma idealnego modelu danego fragmentu rzeczywistości.
- Często istotnie różne modele opisują ten sam fragment rzeczywistości opierając się na różnych założeniach.
- Modele się poprawia lub nawet odrzuca.

Zasada falsyfikowalności (Karl Popper (1902-1994))

Każdy model lub teoria naukowa w obrębie nauk przyrodniczych winny być tak sformułowane, aby **choćby teoretycznie możliwe** było ich obalenie, czyli falsyfikacja.

Jest to bardzo ważny postulat, ale w praktyce często trudno go spełnić.

Modele populacyjne

Zajmiemy się najpierw podstawowymi modelami opisującymi zmiany zagęszczeń populacji.

Przez **zagęszczenie** populacji rozumie się liczbę osobników występujących na obszarze o jednostkowej powierzchni (np. 1 km^2 , 1 m^2) lub w obszarze o jednostkowej objętości (np. 1 m^3 , 1 litr).

Zagęszczenie jest miarą liczebności populacji w danym momencie.

Zagęszczenie jest dobrze charakteryzuje populację, o ile rozmieszczenie osobników w przestrzeni jest względnie **równomierne**.

Dzieje się tak wtedy, gdy ewentualny lokalny wzrost lub spadek liczebności jest szybko wyrównywany na skutek przemieszczania się osobników.

Definicja

Przez **stan populacji** w danej chwili rozumiemy jej zagęszczenie w danej chwili.

Rozważmy najpierw modele z **czasem dyskretnym**

- Interesuje nas stan populacji tylko w wyróżnionych momentach (np. co roku pod koniec sezonu lęgowego) $t_0, t_1, t_2, \dots, t_k \dots$ w równych odstępach czasowych wyznaczonych przez **krok czasowy** Δt , czyli dla dowolnego $k \geq 1$,

$$t_{k+1} - t_k = \Delta t$$

- dla uproszczenia pomijamy strukturę populacji, np. strukturę wieku, płci etc.,

Oznaczając przez N_k stan populacji w momencie t_k (np. w sezonie k) i przez N_0 stan początkowy, otrzymujemy zależność postaci

$$N_{k+1} = N_k + F(N_k),$$

Funkcja $F = F(N)$ opisuje przyrost populacji uwzględniając takie czynniki jak: śmiertelność, reprodukcję, imigrację i emigrację. Przyjmując założenie, że reprodukcja i śmiertelność nie zmieniają się z sezonu na sezon i nie zależą od zagęszczenia populacji w danym sezonie otrzymamy w podstawowym modelu liniowym opisującym zmiany liczebności populacji mamy

$$F(N) = RN = (R_b - R_d)N,$$

gdzie

- R_b — **współczynnik reprodukcji** określony jako liczba potomków, które przeżyły sezon, wydanych przez jednego osobnika,
- R_d — **współczynnik śmiertelności** określający odsetek osobników, które nie przeżyły sezonu.

Porównajmy ten ogólny schemat z rozważanym na wcześniejszych zajęciach modelem wzrostu populacji ptaków

$$N_{k+1} = s(1 + r_0)N_k = N_k + sr_0N_k - (1 - s)N_k,$$

gdzie r_0 to liczba potomków per capita, a s przeżywalność.

Przez R oznaczymy **współczynnik wzrostu populacji *per capita*** (czyli na jednego osobnika) będący różnicą pomiędzy R_b i śmiertelnością R_d .

Po przekształceniu widzimy, że współczynnik wzrostu populacji można wyrazić za pomocą stanów populacji jako

$$R = \frac{N_{k+1} - N_k}{N_k}, \quad \text{dla każdego } k \geq 1.$$

Powyższy model można łatwo rozszerzyć uwzględniając stałe w czasie tempo imigracji, opisywane przez współczynnik R_i oraz tempo emigracji ze współczynnikiem R_e . Wtedy

$$R = R_b - R_d + R_i - R_e.$$

Jeśli reprodukcja przebiega w sposób ciągły, tzn. osobniki młode przychodzą na świat przez cały rok, należy raczej przyjąć model z czasem ciągłym. Prowadzi to do wykorzystania **równań różniczkowych**.

Nie jest to jednak reguła, także wtedy można posłużyć się modelem z czasem dyskretnym, w którym współczynniki R_b i R_d są wielkościami uśrednionymi po całym kroku czasowym Δt .

Nie ma niestety prostej zależności między modelami z czasem ciągłym i modelami z czasem dyskretnym, a niekiedy różnice w przewidywaniach otrzymanych za pomocą obu modeli na długich przedziałach czasu mogą być znaczne.

Model z czasem ciągłym

Przez $N(t)$ oznaczmy zagęszczenie populacji w chwili t .

Ustalmy małe $\Delta t > 0$. Przyrost populacji *per capita* w czasie Δ jest proporcjonalny do czasu Δt ze współczynnikiem proporcjonalności r , czyli

$$\frac{N(t + \Delta t) - N(t)}{N(t)} = r \cdot \Delta t.$$

Teraz dzielimy stronami przez Δt

$$\frac{\frac{N(t+\Delta t) - N(t)}{\Delta t}}{N(t)} = r \xrightarrow{\Delta t \rightarrow 0} \frac{\frac{dN(t)}{dt}}{N(t)} = r \implies \frac{dN(t)}{dt} = r N(t)$$

Przechodząc z Δt po lewej stronie otrzymujemy pochodną funkcji $N(t)$ w chwili t i ostatecznie otrzymujemy *równanie różniczkowe*.

Model z czasem ciągłym

Przez $N(t)$ oznaczmy zagęszczenie populacji w chwili t .

Ustalmy małe $\Delta t > 0$. Przyrost populacji *per capita* w czasie Δ jest proporcjonalny do czasu Δt ze współczynnikiem proporcjonalności r , czyli

$$\frac{N(t + \Delta t) - N(t)}{N(t)} = r \cdot \Delta t.$$

Teraz dzielimy stronami przez Δt

$$\frac{\frac{N(t+\Delta t) - N(t)}{\Delta t}}{N(t)} = r \xrightarrow{\Delta t \rightarrow 0} \frac{\frac{dN(t)}{dt}}{N(t)} = r \implies \frac{dN(t)}{dt} = r N(t)$$

Przechodząc z Δt po lewej stronie otrzymujemy pochodną funkcji $N(t)$ w chwili t i ostatecznie otrzymujemy *równanie różniczkowe*.

Model z czasem ciągłym

Przez $N(t)$ oznaczmy zagęszczenie populacji w chwili t .

Ustalmy małe $\Delta t > 0$. Przyrost populacji *per capita* w czasie Δ jest proporcjonalny do czasu Δt ze współczynnikiem proporcjonalności r , czyli

$$\frac{N(t + \Delta t) - N(t)}{N(t)} = r \cdot \Delta t.$$

Teraz dzielimy stronami przez Δt

$$\frac{\frac{N(t+\Delta t) - N(t)}{\Delta t}}{N(t)} = r \xrightarrow{\Delta t \rightarrow 0} \frac{\frac{dN(t)}{dt}}{N(t)} = r \implies \frac{dN(t)}{dt} = r N(t)$$

Przechodząc z Δt po lewej stronie otrzymujemy pochodną funkcji $N(t)$ w chwili t i ostatecznie otrzymujemy *równanie różniczkowe*.

Model z czasem ciągłym

Przez $N(t)$ oznaczmy zagęszczenie populacji w chwili t .

Ustalmy małe $\Delta t > 0$. Przyrost populacji *per capita* w czasie Δ jest proporcjonalny do czasu Δt ze współczynnikiem proporcjonalności r , czyli

$$\frac{N(t + \Delta t) - N(t)}{N(t)} = r \cdot \Delta t.$$

Teraz dzielimy stronami przez Δt

$$\frac{\frac{N(t+\Delta t)-N(t)}{\Delta t}}{N(t)} = r \xrightarrow{\Delta t \rightarrow 0} \frac{\frac{dN(t)}{dt}}{N(t)} = r \implies \frac{dN(t)}{dt} = r N(t)$$

Przechodząc z Δt po lewej stronie otrzymujemy pochodną funkcji $N(t)$ w chwili t i ostatecznie otrzymujemy *równanie różniczkowe*.

Model z czasem ciągłym

Przez $N(t)$ oznaczmy zagęszczenie populacji w chwili t .

Ustalmy małe $\Delta t > 0$. Przyrost populacji *per capita* w czasie Δ jest proporcjonalny do czasu Δt ze współczynnikiem proporcjonalności r , czyli

$$\frac{N(t + \Delta t) - N(t)}{N(t)} = r \cdot \Delta t.$$

Teraz dzielimy stronami przez Δt

$$\frac{\frac{N(t+\Delta t)-N(t)}{\Delta t}}{N(t)} = r \xrightarrow{\Delta t \rightarrow 0} \frac{\frac{dN(t)}{dt}}{N(t)} = r \implies \frac{dN(t)}{dt} = r N(t)$$

Przechodząc z Δt po lewej stronie otrzymujemy pochodną funkcji $N(t)$ w chwili t i ostatecznie otrzymujemy **równanie różniczkowe**.

Model z czasem ciągłym

Czyli przyrost populacji w chwili t jest opisywany przez równanie różniczkowe:

$$\dot{N}(t) = \frac{dN(t)}{dt} = rN(t). \quad (*)$$

Przyjmujemy też, że zagęszczenie populacji w jakiejś ustalonej chwili $t = t_0$ równe jest

$$N(t_0) = N_0. \quad (**)$$

W ten sposób zadajemy **warunek początkowy** dla równania różniczkowego (*).

W równaniu różniczkowym (*) z warunkiem początkowym (**) szukamy **funkcji**, która będzie określona na odcinku zawierającym t_0 , takiej że spełnia równanie (*) oraz jej wartość w t_0 jest równa N_0 .

Równanie Malthusa

Równanie

$$\dot{N}(t) = \frac{dN(t)}{dt} = rN(t), \quad N(t_0) = N_0. \quad (*)$$

jest podstawowym **liniowym** równaniem różniczkowym opisującym zmiany zagęszczeń populacji w czasie ciągłym.

Równanie to zwane jest równaniem Malthusa (Thomas Malthus — angielski ekonomista — żył na przełomie XVIII w. i XIX w. i prowadził badania z pogranicza ekonomii i socjologii).

Współczynnik r nazywany jest **współczynnikiem wzrostu populacji** lub **współczynnikiem wewnętrznego tempa wzrostu populacji**.

Uwaga

Współczynniki wzrostu: R w modelu z czasem dyskretnym i r w modelu z czasem ciągłym, choć pełnią podobną rolę **nie są sobie równe**.

Współczynnik r może być dodatni bądź ujemny i podobnie jak w przypadku modelu z czasem dyskretnym przedstawia się jako różnica

$$r = r_b - r_d$$

gdzie r_b uwzględnia jedynie liczbę potomków przychodzących na świat w jednostce czasu i pochodzących średnio od jednego osobnika, a r_d liczbę „zgonów” przypadających w jednostce czasu na jednego osobnika.

Współczynnik r wyraża się w jednostkach [1/czas].

Przykład

Jeśli zagęszczenie wyrażamy w jednostkach [liczba osobników/(litr)], a czas w godzinach, to po lewej stronie równania mamy jednostki [liczba osobników/(godz. \times litr)]. Stąd wynika, że r wyraża się w jednostce [1/(godz.)]

Rozwiązanie równania Malthusa

Dzielimy obie strony równania Malthusa przez $N(t)$ i całkujemy od t_0 do T . Dostajemy wtedy

$$\int_{t_0}^T \frac{\dot{N}(t)}{N(t)} dt = \int_{t_0}^T \frac{d}{dt}(\ln N(t)) dt = \int_{t_0}^T r dt = r(T - t_0),$$

czyli

$$\ln N(T) - \ln N_0 = r(T - t_0),$$

i ostatecznie, wracając do typowego oznaczenia zmiennej czasu literą t , otrzymujemy rozwiązanie

$$N(t) = N_0 e^{r(t-t_0)} \quad \text{dla } t \in \mathbb{R}.$$

Równania różniczkowe pojawiają się we wszystkich naukach przyrodniczych.

W biologii modele oparte o równania różniczkowe pojawiają się na wszystkich poziomach organizacji materii biologicznej od subkomórkowego po ekologiczny.

Równanie różniczkowe przedstawia sposób w jaki stan danego procesu w danej chwili wpływa na tempo jego zmian (wyrażone za pomocą pochodnej).

Na przykład: w równaniu Malthusa tempo zmiany zagęszczenia jest w każdej chwili wprost proporcjonalne do zagęszczenia populacji. Im większa populacja tym, proporcjonalnie, większa (lub tym mniejsza) szybkość zmian jej zagęszczeń.

Podstawowe własności równań różniczkowych

$$\dot{x}(t) = f(x(t)), \quad (*)$$

- Rozwiązaniem równania (*) jest funkcja, która po wstawieniu za x spełnia to równanie. Najczęściej do równania różniczkowego dołącza się warunek początkowy np. $x(t_0) = x_0$. Wtedy otrzymujemy tzw. zagadnienie początkowe polegające na określeniu takiej funkcji, która spełnia równanie różniczkowe i w chwili $t = t_0$ przyjmuje wartość x_0 .
- Przeważnie nie umiemy znaleźć wzoru (formuły) na rozwiązanie równania (*), oczywiście poza specjalnymi klasami równań takimi jak równania liniowe i niektóre równania nieliniowe .

Stabilność punktu stacjonarnego

Definicja

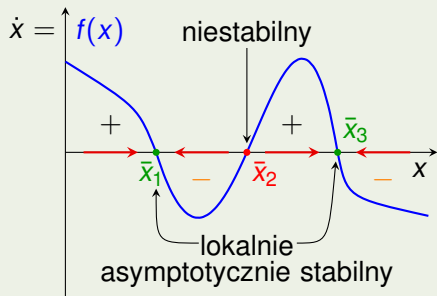
Punktem stacjonarnym równania (\star) nazywamy taki punkt \bar{x} , że jeśli $x(0) = \bar{x}$, to wówczas $x(t) = \bar{x}$ dla wszystkich $t \geq 0$. Zatem punkt stacjonarny spełnia równanie $0 = f(\bar{x})$. Rzeczywiście, jeśli $x(t) = \bar{x}$ dla wszystkich $t \in \mathbb{R}$, to $\dot{x}(t) = 0 = f(x(t))$.

Definicja

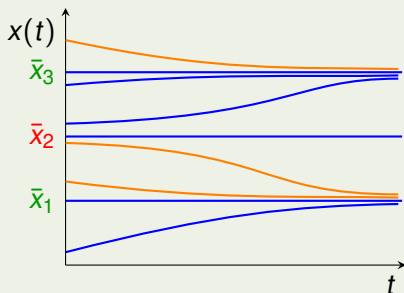
Punkt stacjonarny równania różniczkowego (\star) jest **lokalnie asymptotycznie stabilny** jeśli istnieje takie otoczenie tego punktu, że rozwiązanie startujące z dowolnego punktu z tego otoczenia zbiega w granicy do tego stanu stacjonarnego.

Jak określić zachowanie rozwiązań na podstawie wykresu funkcji f ?

Portret fazowy



Wykres rozwiązań



- $x(0) \in (0, \bar{x}_1) \implies x(t)$ rośnie do \bar{x}_1
- $x(0) \in (\bar{x}_1, \bar{x}_2) \implies x(t)$ **maleje** do \bar{x}_1
- $x(0) \in (\bar{x}_2, \bar{x}_3) \implies x(t)$ rośnie do \bar{x}_3
- $x(0) > \bar{x}_3 \implies x(t)$ **maleje** do \bar{x}_3

Jak analitycznie zbadać stabilność punktów stacjonarnych?

- Jeżeli $f'(\bar{x}) < 0$, to wówczas punkt stacjonarny \bar{x} jest lokalnie asymptotycznie stabilny;
- Jeżeli $f'(\bar{x}) > 0$, to wówczas punkt stacjonarny \bar{x} jest niestabilny;
- Jeżeli $f'(\bar{x}) = 0$, to metoda nie rozstrzyga stabilności punktu stacjonarnego.

Równanie logistyczne

Tak jak w przypadku modelu z czasem dyskretnym, liniowy model wzrostu (model Mathusa) nie uwzględnia ograniczeń środowiskowych (np. ograniczonej ilości pokarmu).

Uwzględnienie czynników zagęszczeniozależnych

Przyjmijmy, że współczynnik wzrostu populacji jest proporcjonalny do zasobów środowiskowych, które maleją wraz ze wzrostem liczebności populacji, czyli założymy, że $r = \tilde{r}(N)$ jest **funkcją malejącą**.

Wówczas przyrost *per capita* wynosi

$$\frac{\dot{N}(t)}{N(t)} = \tilde{r}(N(t)) \implies \dot{N}(t) = \tilde{r}(N(t))N(t).$$

Przyjmujemy $\tilde{r}(N) = r(1 - N/K)$ i otrzymujemy

równanie logistyczne (lub równanie Verhulsta)

$$\dot{N}(t) = r N(t) \left(1 - \frac{N(t)}{K}\right).$$

Właściwości równania logistycznego

Równanie logistyczne z czasem ciągłym (Verhulsta)

$$\dot{N}(t) = r N(t) \left(1 - \frac{N(t)}{K} \right).$$

Postać tę zaproponował ok. 1840 roku Verhulst (Pierre F. Verhulst (1804-1849)). Parametr K nosi nazwę **pojemności środowiska** i równy jest zagęszczeniu populacji, przy którym tempo wzrostu populacji spada do zera. Jest to zarazem maksymalne zagęszczenie populacji wyznaczone przez wielkość zasobów siedliska.

Alternatywne wyprowadzenie polega na uzupełnieniu równania Malthusa o składnik aN wyrażający wzrost śmiertelności proporcjonalny do zagęszczenia :

$$\dot{N}(t) = (r - aN(t))N(t) = rN(t)\left(1 - \frac{a}{r}N\right)$$

wtedy $K = \frac{r}{a}$.

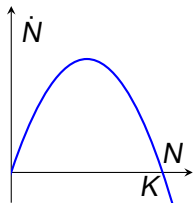
Własności równania logistycznego

Równanie logistyczne z czasem ciągłym (Verhulsta)

$$\dot{N}(t) = r N(t) \left(1 - \frac{N(t)}{K} \right).$$

Gdy wielkość N jest stosunkowo mała, to wyrażenie w nawiasie jest bliskie 1 i wtedy wzrost populacji jest porównywalny do wzrostu w równaniu Malthusa (parametr r pełni tu tę samą rolę co w równaniu Malthusa).

Gdy N jest bliskie K , to wyrażenie w nawiasie zbliża się do zera całkowicie redukując tempo wzrostu populacji.



Jeśli $N_0 \in (0, K)$, to rozwiązanie jest funkcją rosnącą dążącą do K .

Jeśli $N_0 > K$, to rozwiązanie jest funkcją malejącą do K .

Jeśli $N_0 = 0$ lub $N_0 = K$, to funkcje stałe w czasie $N(t) = 0$ dla $t \geq 0$ oraz $N(t) = K$ dla $t \geq 0$ są rozwiązaniami stacjonarnymi.

Własności równania logistycznego

Można znaleźć wzór na rozwiązanie równania logistycznego!

$$N(t) = \frac{1}{K^{-1} + e^{-rt}(N_0^{-1} - K^{-1})} \quad \text{dla } t \in \mathbb{R}.$$

Dla wszystkich $N_0 > 0$ mamy $\lim_{t \rightarrow +\infty} N(t) = K$, zatem stan stacjonarny $\bar{N} = K$ jest **globalnie** asymptotycznie stabilny.

Funkcja $N(t)$ wzrasta/maleje asymptotycznie do K (to znaczy zbliża się do K , gdy $t \rightarrow +\infty$, nigdy nie osiągając tego stanu!). Warto wspomnieć o jeszcze jednej ciekawej własności.

Zróżniczkujmy równanie logistyczne względem czasu.

$$\ddot{N}(t) = r\dot{N}(t) \left(1 - \frac{N(t)}{K}\right) - \frac{r}{K}N(t)\dot{N}(t) = r\dot{N}(t) \left(1 - \frac{2}{K}N(t)\right).$$

Druga pochodna rozwiązania $\ddot{N}(t)$ zeruje się zawsze w takiej chwili t , że

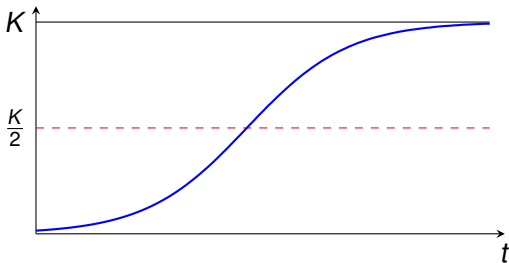
$$N(t) = \frac{K}{2}.$$

Krzywa logistyczna

Rozwiązanie przyjmujące wartości w przedziale $(0, K)$ jest funkcją rosnącą. Zatem druga pochodna zeruje się tylko gdy $N_0 \in (0, \frac{K}{2})$.

Każde rozwiązanie startujące z przedziału $(0, \frac{K}{2})$ rośnie w sposób wypukły aż do punktu przegięcia ($\ddot{N} = 0$), gdy zagęszczenie osiąga połowę pojemności siedliska. Potem rośnie coraz wolniej (jest funkcją wklęsłą).

Wykres tej funkcji ma charakterystyczny kształt przypominający rozciągniętą literę S. Tę krzywą nazywa się **krzywą logistyczną**.



Krzywa logistyczna jest bardzo często wykorzystywana do modelowania efektu wysycenia, matematycznie ujętego jako asymptotyczne zbliżanie się do granicy.

Rozwiązania startujące z przedziału $(\frac{K}{2}, K)$ są funkcjami wklęsłymi (nie mają punktów przegięcia).

Rozwiązanie z warunkiem początkowym $N_0 \in (K, +\infty)$ może odpowiadać pojawieniu się (np. introdukcji) jakiegoś gatunku, np. na wyspie, w sytuacji gdy liczebność początkowa przekracza pojemność siedliska.

Wtedy zagęszczenie populacji maleje w czasie zbiegając asymptotycznie do K .

Jakie by nie było początkowe zagęszczenie populacji (na ogół trudne do precyzyjnego ustalenia), zbiega ono w czasie do jedynego zagęszczenia K , co w praktyce oznacza, że prędzej czy później zagęszczenie populacji osiągnie wartości dowolnie bliskie K .