STATYSTYKA OPISOWA

Dr Alina Gleska

Instytut Matematyki WE PP

28 września 2018

SZEREGIEM CZASOWYM nazywamy tablicę, która zawiera ciąg wartości cechy uporządkowanych według kolejnych jednostek czasu. Zatem szereg czasowy przedstawia rozwój wartości cechy w czasie, w wybranych okresach lub momentach.
SZEREGI CZASOWE OKRESÓW zawierają informacje o rozmiarach zjawiska w krótszych lub dłuższych okresach (kwartał, rok).
SZEREGI CZASOWE MOMENTÓW ujmują wielkość zjawiska w danym momencie, najczęściej na początku lub końcu miesiąca (np. stan zapasów w przedsiębiorstwie na 31 grudnia danego roku).

Kształtowanie się określonego zjawiska w czasie może być przedstawione za pomocą modeli składowych szeregu czasowego. W praktyce korzysta się z dwóch modeli: addytywnego lub multiplikatywnego.

Model addytywny to $y_t = f(t) + s_i(t) + e(t)$,
model multiplikatywny to $y_t = f(t)s_i(t)e(t)$, gdzie

- y_t – poziom zjawiska w okresie t,
- $f(t)$ – poziom zjawiska w okresie t oszacowany na podstawie funkcji tendencji rozwojowej (czyli trendu),
- $s_i(t)$ – poziom zjawiska w okresie t oszacowany na podstawie funkcji sezonowości, przy czym $i = 1, 2, \ldots, d$ oznacza liczbę podokresów cyklu wahań,
- $e(t)$ – składnik resztowy modelu, zwany składnikiem nieregularnym.

FUNKCJE TENDENCJI ROZWOJOWYCH $f(t)$ przedstawiają rozwój zjawiska w czasie, jego regularne i systematyczne zmiany w ciągu długiego okresu (np. tendencję rozwojową budownictwa mieszkaniowego w Polsce w latach 2000-2016).

ZMIANY SEZONOWE $s_i(t)$ są to zmiany, które powtarzają się regularnie w tym samym okresie każdego roku. Przykładowo, na wahanie zjawisk gospodarczych w rolnictwie mają wpływ różne czynniki natury biologiczno-technicznej oraz klimatyczne, które powtarzają się w określonych porach roku (sezonach).

ZMIANY PRZYPADKOWE $e(t)$, zwane nieregularnymi lub losowymi, wynikają z oddziaływania na dane zjawisko czynników nie dających się przewidzieć (np. pożary, powodzie itp.).

Do oceny dynamiki zjawisk można wykorzystać szereg różnych miar dynamiki. Najważniejsze miary dynamiki:

a) przyrosty absolutne,
b) indeksy indywidualne,
c) stopa zmian,
d) średnie ruchome,
e) indeksy agregatowe dla wielkości absolutnych,
f) indeksy agregatowe dla wielkości stosunkowych,
g) oceny parametrów strukturalnych modeli tendencji rozwojowej (trendu),
h) wskaźniki sezonowości.

Miary dynamiki wymienione w punktach a)-d) oraz g) i h) służą do opisu zmian zachodzących w wartościach jednej cechy (zjawisk jednorodnych), natomiast indeksy agregatowe dla wielkości absolutnych i stosunkowych wykorzystuje się do charakteryzowania zmian występujących w rozwoju więcej niż jednego zjawiska jednocześnie.

PRZYROSTY ABSOLUTNE ujmuje się jako:

- jednopodstawowe – biorąc stałą podstawę porównań y_c:

$$y_1 - y_c, \ y_2 - y_c, \ y_3 - y_c, \dots, y_{n-1} - y_c, \ y_n - y_c,$$

  gdzie za podstawę przyjmujemy $y_c = y_t$ dla pewnego $t \in \{1, 2, \dots, n\}$, albo średnią wieloletnią z pewnego przedziału czasowego $y_c = \overline{y}$;

- łańcuchowe – kiedy za podstawę porównań przyjmuje się wartość cechy z okresu bezpośrednio poprzedzającego okres badany:

$$y_2 - y_1, \ y_3 - y_2, \dots, y_{n-1} - y_{n-2}, \ y_n - y_{n-1}.$$

INDEKSAMI DYNAMIKI nazywamy miary określające względne zmiany w wielkościach badanego zjawiska odnoszącego się do czasu lub miejsca. Indeksy otrzymujemy z podzielenia wartości cechy w badanej jednostce czasu lub miejsca.
INDEKSY INDYWIDUALNE ujmuje się jako przyrosty względne, indeksy jednopodstawowe i łańcuchowe.
PRZYROSTY WZGLĘDNE (wskaźniki tempa przyrostu) wyrażają stosunek przyrostu absolutnego $(y_t - y_c)$ do poziomu zjawiska z okresu podstawowego (bazowego) y_c:

$$P_{t/c} = \frac{y_t - y_c}{y_c} \cdot 100\%$$

lub do poziomu zjawiska z okresu bezpośrednio poprzedzającego okres badany y_{t-1}:

$$P_{t/t-1} = \frac{y_t - y_{t-1}}{y_{t-1}} \cdot 100\% \quad \text{dla} \quad t > 2.$$

Wartości przyrostów względnych większe od 0 świadczą o wzroście poziomu zjawiska w okresie badanym w porównaniu z okresem podstawowym, a mniejsze od 0 świadczą o spadku poziomu danego zjawiska.

INDEKSY INDYWIDUALNE mogą być obliczane jako

$$\text{jednopodstawowe:} \quad I_{t/c} = \frac{y_t}{y_c} \cdot 100\%$$

lub

$$\text{łańcuchowe:} \quad I_{t/t-1} = \frac{y_t}{y_{t-1}} \cdot 100\% \quad \text{dla } t > 2.$$

Wartości indeksów większe od 100% świadczą o wzroście poziomu zjawiska w okresie badanym w porównaniu z okresem podstawowym, a gdy są mniejsze od 100% – świadczą o spadku danego zjawiska.

Indeksy indywidualne mają następujące właściwości:

- iloraz dwóch indeksów jednopodstawowych o tej samej podstawie c dla poziomu zjawiska z różnych okresów k i l jest indeksem jednopodstawowym:

$$\frac{I_{k/c}}{I_{l/c}} = \frac{y_k}{y_c} : \frac{y_l}{y_c} = \frac{y_k}{y_l} = I_{k/l};$$

- iloraz dwóch indeksów jednopodstawowych o tej samej podstawie c dla poziomu zjawiska z kolejnych dwóch okresów k i $k-1$ jest indeksem łańcuchowym. Zamiana indeksów jednopodstawowych o podstawie c w łańcuchowy przebiega następująco:

$$\frac{I_{k/c}}{I_{k-1/c}} = \frac{y_k}{y_c} : \frac{y_{k-1}}{y_c} = \frac{y_k}{y_{k-1}} = I_{k/k-1};$$

- iloczyn lub odwrotność iloczynu kolejnych indeksów łańcuchowych jest indeksem jednopodstawowym:
  - gdy $k > c$:

$$I_{c+1/c} \cdot I_{c+2/c+1} \cdots I_{k/k-1} = \frac{y_{c+1}}{y_c} \cdot \frac{y_{c+2}}{y_{c+1}} \cdots \frac{y_k}{y_{k-1}} = \frac{y_k}{y_c} = I_{k/c},$$

  - gdy $k < c$:

$$\frac{1}{I_{k+1/k} \cdot I_{k+2/k+1} \cdots I_{c/c-1}} = \frac{1}{\frac{y_{k+1}}{y_k} \cdot \frac{y_{k+2}}{y_{k+1}} \cdots \frac{y_c}{y_{c-1}}} = \frac{1}{\frac{y_c}{y_k}} = \frac{y_k}{y_c} = I_{k/c};$$

- przyrost względny powiększony o jeden jest indeksem łańcuchowym, czyli indeks łańcuchowy pomniejszony o jeden jest przyrostem względnym:

$$
\begin{aligned}
P_{t/t-1} + 1 &= \frac{y_t - y_{t-1}}{y_{t-1}} + 1 = \frac{y_t - y_{t-1} + y_{t-1}}{y_{t-1}} = \frac{y_t}{y_{t-1}} = I_{t/t-1}, \\
I_{t/t-1} - 1 &= \frac{y_t}{y_{t-1}} - 1 = \frac{y_t - y_{t-1}}{y_{t-1}} = P_{t/t-1}.
\end{aligned}
$$

Średnia ruchoma jest wykorzystywana do wygładzania szeregu czasowego, czyli oczyszczania ze zmian przypadkowych. Idea wyrównywania szeregu czasowego za pomocą średnich ruchomych polega na zastąpieniu pierwotnych wartości cechy średnimi arytmetycznymi, obliczanymi sekwencyjnie dla wybranej liczby obserwacji. Wyznaczone wartości średnie przyporządkowuje się na ogół środkowym obserwacjom, na podstawie których były obliczane średnie.

Używając średniej ruchomej do badania tendencji rozwojowej przyjmuje się, że wartość cechy w badanym momencie lub okresie będzie równa średniej arytmetycznej z nieparzystej liczby wyrazów szeregu czasowego, czyli

$$\overline{y_t} = \frac{1}{k} \sum_{i=-s}^{s} y_{t+i},$$

gdzie $k = 2s + 1$ jest nazywana stałą wygładzania ($k = 3, 5, 7, \ldots$), a $s < t \leq n - s$.

Do obliczenia średniej na podstawie parzystej liczby wyrazów (np. dla kwartałów lub miesięcy) używamy średniej ruchomej scentrowanej:

$$\overline{y_t} = \frac{1}{k}(0.5y_{t-s} + \sum_{i=-s+1}^{s-1} y_{t+i} + 0.5y_{t+s})$$

gdzie $k = 2s$, a $s < t \leq n - s$.

Tabela średnich ruchomych

Miarę dynamiki zwaną stopą zmian wykorzystuje się do charakteryzowania średniego tempa zmian wartości cechy za cały badany okres. Stopę zmian można obliczać w dwojaki sposób, na podstawie:

- dwóch skrajnych wyrazów szeregu czasowego,
- wszystkich wyrazów szeregu czasowego.

Stosując pierwszy sposób oblicza się najpierw średnią geometryczną z indeksów łańcuchowych y_t/y_{t-1} $(t = 2, 3, \ldots, n)$:

$$\bar{I}_{t/t-1} = \sqrt[n-1]{\frac{y_2}{y_1} \cdot \frac{y_3}{y_2} \cdots \frac{y_n}{y_{n-1}}} = \sqrt[n-1]{\frac{y_n}{y_1}} = \left(\frac{y_n}{y_1}\right)^{\frac{1}{n-1}}.$$

Po zlogarytmowaniu obu stron równania otrzymujemy:

$$ln\bar{I}_{t/t-1} = ln\left(\frac{y_n}{y_1}\right)^{\frac{1}{n-1}} = \frac{1}{n-1} ln\left(\frac{y_n}{y_1}\right) = C,$$

gdzie C – konkretna wartość.

Wówczas:

$$\bar{I}_{t/t-1} = e^{C} = e^{\frac{1}{n-1} ln\left(\frac{y_n}{y_1}\right)}.$$

Po obliczeniu $\bar{I}_{t/t-1}$ wyznaczamy średniookresowe tempo zmian jako:

$$r_g = \left(\bar{I}_{t/t-1} - 1\right) \cdot 100\%.$$

Obliczanie stopy zmian na podstawie dwóch skrajnych wartości cechy jest wskazane w przypadku, gdy badane zjawisko wykazuje jednoznaczną tendencję rozwojową, wzrostową albo spadkową.

Jeżeli w rozwoju zjawiska występują wyraźne odchylenia od tendencji wzrostowej albo spadkowej, to bardziej poprawne jest obliczanie średniego tempa zmian na podstawie wszystkich wartości cechy za pomocą wzoru:

$$r_g = \frac{-3m + \sqrt{9m^2 + 24m(n-1)\left(\frac{1}{y_1}\sum_{t=1}^{n} y_t - n\right)}}{2m(n-1)} \cdot 100\%,$$

gdzie $m = n(n+1)$.
Jeżeli stopa zmian jest większa od 0, to badane zjawisko wykazuje tendencję wzrostową. W przypadku, gdy stopa zmian jest mniejsza od 0 mamy do czynienia z tendencją spadkową. Gdy stopa zmian jest równa 0, to występuje stagnacja w rozwoju zjawiska.

W analizach finansowych stopa zmian jest określana mianem stopy procentowej. Posługujemy się nią w obliczaniu przyszłej wartości pieniądza y_n, po n latach, przy założeniu, że zainwestowany kapitał początkowy wynosi y_0 (przy stałej stopie procentowej).
Wartość y_n wyznacza się w zależności od różnych form i częstotliwości kapitalizacji odsetek, stosując:

- kapitalizację zwykłą (tylko kapitał początkowy przynosi dochód)

$$y_n = y_0(1 + n \cdot r_g),$$

- roczną kapitalizację odsetek:

$$y_n = y_0(1 + r_g)^n,$$

- wielokrotną w ciągu roku kapitalizację odsetek:

$$y_n = y_0(1 + \frac{r_g}{m})^{n \cdot m},$$

gdzie y_n oznacza przyszłą wartość pieniądza, y_0 – wartość zainwestowanego kapitału (kapitał początkowy), r_g – stałą stopę procentową, n – liczbę lat kapitalizacji odsetek, a m – liczbę okresów w ciągu roku.