



**Joanna Małgorzata Landmesser**

# **EKONOMETRIA I PROGNOZOWANIE PROCESÓW EKONOMICZNYCH**

## **Literatura podstawowa:**

1. Borkowski B., Dudek H., Szczesny W.: „Ekonometria. Wybrane zagadnienia”, PWN, Warszawa 2003.
2. Gruszczyński M., Kluza S., Winek D.: „Ekonometria”, WSHiFM, Warszawa 2003.
3. Kufel T.: „Ekonometria. Rozwiązywanie problemów z wykorzystaniem programu gretl”, PWN, Warszawa 2004.

## **Literatura uzupełniająca:**

1. Maddala G. S.: „Ekonometria”, PWN, Warszawa 2006.
2. Aleksander Welfe: „Ekonometria”, PWE, Warszawa 1995.
3. Nowak E.: „Zarys metod ekonometrii. Zbiór zadań”, PWN, Warszawa 2002.
- Zeliaś A., Pawełek B., Wanat S.: „Prognozowanie ekonomiczne. Teoria, przykłady, zadania”, PWN, Warszawa 2003.

## Spis treści

<b>I. Dobór zmiennych objaśniających do modelu ekonometrycznego</b>	<b>3</b>
1. Zasady doboru zmiennych	3
2. Metoda eliminacji zmiennych quasi-stałych	3
3. Metoda analizy macierzy współczynników korelacji	3
4. Metoda Hellwiga doboru zmiennych do modelu	5
<b>II. Klasyczny model regresji</b>	<b>6</b>
1. Zapis modelu	6
2. Założenia modelu liniowej regresji wielu zmiennych	6
3. Klasyczna metoda najmniejszych kwadratów	7
4. Interpretacja współczynników równania	9
<b>III. Weryfikacja modelu</b>	<b>10</b>
1. Ocena jakości modelu	10
2. Badanie istotności zmiennych objaśniających modelu	12
3. Weryfikacja reszt	16
<b>IV. Analiza rozwoju zjawisk w czasie</b>	<b>21</b>
1. Szeregi czasowe	21
2. Modele trendu	22
3. Wygładzanie szeregu czasowego metodą średniej ruchomej	24
4. Stacjonarność procesów stochastycznych	25
5. Modele AR, MA, ARMA i ARIMA	27
6. Inne modele ze zmiennymi opóźnionymi	29
7. Kointegracja szeregów czasowych	30
<b>V. Prognozowanie ekonometryczne</b>	<b>31</b>
1. Prognoza	31
2. Błędy prognozy	32
3. Prognozowanie na podstawie modeli tendencji rozwojowej	33
4. Inne metody prognozowania	34
<b>VI. Modele ze zmiennymi jakościowymi</b>	<b>36</b>
1. Zmienne jakościowe w charakterze zmiennych objaśniających	36
2. Wykorzystanie zmiennych jakościowych do analizy danych sezonowych	37
3. Zmienne jakościowe w charakterze zmiennych objaśnianych	38
<b>VII. Modele nieliniowe</b>	<b>40</b>
1. Klasyfikacja modeli nieliniowych	40
2. Typy modeli nieliniowych stosowane w badaniach ekonomicznych	40
3. Metody estymacji modeli nieliniowych	42
<b>VIII. Ekonometryczna analiza produkcji i popytu</b>	<b>44</b>
1. Ekonometryczna analiza produkcji	44
2. Ekonometryczna analiza popytu	47
<b>IX. Wielorównaniowe modele ekonometryczne</b>	<b>49</b>
1. Klasyfikacja zmiennych	49
2. Postać strukturalna i zredukowana modelu	50
3. Klasyfikacja modeli wielorównaniowych	51
4. Estymacja parametrów modeli wielorównaniowych	52



# I. DOBÓR ZMIENNYCH OBJAŚNIAJĄCYCH DO MODELU EKONOMETRYCZNEGO

1. Zasady doboru zmiennych
2. Metoda eliminacji zmiennych quasi-stałych
3. Metoda analizy macierzy współczynników korelacji
4. Metoda Hellwiga doboru zmiennych do modelu

## 1. ZASADY DOBORU ZMIENNYCH

Zazwyczaj zestaw zmiennych, które mają wystąpić w modelu w charakterze zmiennych objaśniających nie jest określony jednoznacznie. Wynika to przykładowo z:

- mało precyzyjnej teorii ekonomicznej,
- większej liczebności zbioru potencjalnych zmiennych objaśniających niż liczba dostępnych obserwacji.

Metody doboru zmiennych objaśniających do modelu składają się z dwóch elementów:

1. kryterium, na podstawie którego można porównywać różne zestawy zmiennych,
2. sposobu „przeglądania” zbioru potencjalnych zmiennych objaśniających.

Ad 1. Tu stosuje się mierniki jakości modelu (np. współczynnik determinacji) lub ich zestawy (np. współczynnik determinacji + wymóg statystycznej istotności ocen parametrów).

Ad 2.

- metoda eliminacji a priori (zaczynam od jednej zmiennej i ewentualnie dokładam do modelu kolejne),
- metoda selekcji a posteriori (zaczynam od modelu ze wszystkimi zmiennymi, potem usuwam kolejno najgorsze z nich),
- metoda przeglądu bezpośredniego (badanie własności wszystkich kombinacji zbioru potencjalnych zmiennych objaśniających i wybór najlepszej).

## 2. METODA ELIMINACJI ZMIENNYCH QUASI-STAŁYCH

Miarą poziomu zmienności zmiennej jest współczynnik zmienności:

$$v_i = \frac{S_i}{\bar{x}_i},$$

gdzie  $\bar{x}_i = \frac{1}{m} \sum_{t=1}^m x_{it}$  jest średnią arytmetyczną dla zmiennej  $X_i$ , a  $S_i = \left( \frac{1}{m} \sum_{t=1}^m (x_{it} - \bar{x}_i)^2 \right)^{\frac{1}{2}}$

odchyleniem standardowym zmiennej  $X_i$ .

Po ustaleniu przez badacza wartości krytycznej  $v^*$  przyrównuje się ją do wartości współczynnika zmienności dla danej zmiennej. Jeśli  $v_i \leq v^*$ , to zmienną uznaje się za quasi-stałą i eliminuje ze zbioru potencjalnych zmiennych objaśniających.

## 3. METODA ANALIZY MACIERZY WSPÓŁCZYNNIKÓW KORELACJI

Niech  $Y$  – zmienna objaśniana,  $X_1, \dots, X_m$  – potencjalne zmienne objaśniające („kandydatki”),  $n$  - ilość obserwacji na zmiennych.

Metoda analizy macierzy współczynników korelacji polega na wyborze takich zmiennych objaśniających do modelu, które są silnie skorelowane ze zmienną objaśnianą i jednocześnie słabo skorelowane pomiędzy sobą.

**Przypomnienie:**

Do badania zależności dwóch cech mierzalnych służy **współczynnik korelacji liniowej Pearsona**. Jest to stosunek iloczynów odchyłeń wartości zmiennych od średnich arytmetycznych do iloczynu liczebności zbiorowości i odchyłeń standardowych tych zmiennych:

$$r_{xy} = \frac{\sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{n \sigma_x \sigma_y} = \frac{\text{cov}_{xy}}{\sigma_x \sigma_y} = \frac{\sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

gdzie:  $r_{xy}$  - współczynnik korelacji liniowej Pearsona

$n$  - liczebność próby,

$x_i, y_i$  - wartości zmiennej  $X$  i zmiennej  $Y$ ,

$\bar{x}, \bar{y}$  - średnie arytmetyczne zmiennej  $X$  i zmiennej  $Y$ ,

$\sigma_x, \sigma_y$  - odchylenia standardowe zmiennej  $X$  i zmiennej  $Y$ ,

$$\sigma_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}, \quad \sigma_y = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n}}$$

Współczynnik korelacji liniowej Pearsona przyjmuje wartości z przedziału  $\langle -1, 1 \rangle$ .

W przypadku braku związku między zmiennymi  $r_{xy} = 0$ . W przypadku związku funkcyjnego dodatniego bądź ujemnego (korelacja doskonała) współczynnik przyjmuje wartość 1 lub  $-1$ . Im siła zależności między  $X$  i  $Y$  większa, tym wartość bezwzględna współczynnika jest bliższa jedności.

Punktem wyjścia w metodzie analizy macierzy współczynników korelacji jest znajomość wektora  $\mathbf{R}_0$  i macierzy  $\mathbf{R}$ :

$$\mathbf{R}_0 = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix}, \text{ gdzie } r_i \text{ jest współczynnikiem korelacji pomiędzy zmienną objaśnianą } Y \text{ oraz}$$

potencjalną zmienną objaśniającą  $X_i$ .

$$\mathbf{R} = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1m} \\ r_{21} & 1 & \cdots & r_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ r_{m1} & r_{m2} & \cdots & r_{mm} \end{bmatrix}, \text{ gdzie } r_{ij} \text{ jest współczynnikiem korelacji pomiędzy potencjalnymi}$$

zmiennymi objaśniającymi  $X_i$  oraz  $X_j$ .

Macierz  $\mathbf{R}$  jest symetryczna (tzn.  $r_{ij} = r_{ji}$ ), a na jej diagonalnej znajdują się jedyńki (ponieważ  $r_{ii} = 1$ ).

Przyjmując poziom istotności  $\alpha$  oraz dla  $n-2$  stopni swobody wyznacza się wartość krytyczną współczynnika korelacji:

$$r^* = \left( \frac{(t^*)^2}{(t^*)^2 + n - 2} \right)^{1/2},$$

gdzie  $t^*$  jest wartością statystyki odczytanej z tablic dla rozkładu t-Studenta dla  $\alpha$  i  $n-2$  stopni swobody. Kolejno wykonywane są kroki następującej procedury:

1. Ze zbioru potencjalnych zmiennych objaśniających eliminuje się te zmienne, dla których zachodzi:

$$|r_i| \leq r^*,$$

są to bowiem zmienne nieistotnie skorelowane ze zmienną objaśnianą.



2. Spośród pozostałych zmiennych jako zmienną objaśniającą w modelu powołuje się taką zmienną  $X_h$ , dla której:

$$|r_h| = \max_i \{|r_i|\};$$

zmienną  $X_h$  zwie się nośnikiem największego zasobu informacji o zmiennej objaśnianej.

3. Ze zbioru potencjalnych zmiennych objaśniających eliminuje się te zmienne, dla których zachodzi:

$$|r_{hi}| > r^*,$$

są to bowiem zmienne zbyt silnie skorelowane ze zmienną objaśniającą  $X_h$ , czyli powielające informacje wnoszone do modelu przez tą zmienną.

Powyższą procedurę powtarzamy, aż do momentu wyczerpania zbioru potencjalnych zmiennych objaśniających.

#### 4. METODA HELLWIGA DOBORU ZMIENNYCH DO MODELU

Idea tej metody polega na wyborze optymalnej kombinacji zmiennych objaśniających, tzn. takiej, której elementy będą w największym stopniu skorelowane z zmienną endogeniczną i równocześnie słabo skorelowane między sobą w porównaniu z innymi możliwymi kombinacjami zmiennych.

Nośnikiem informacji o zmiennej endogenicznej jest potencjalna zmienna objaśniająca. Liczba wszystkich możliwych kombinacji potencjalnych zmiennych objaśniających spośród  $m$  zmiennych wynosi:  $C = 2^m - 1$ .

**Pojemnością indywidualną nośnika informacji**, będącą miernikiem wielkości informacji o zmiennej  $Y$  wnoszonej przez zmienną  $X_j$  w  $l$ -tej kombinacji, jest

$$h_{lj} = \frac{r_j^2}{\sum_{i \in I_l} |r_{ij}|} \quad (l = 1, 2, \dots, C; j \in I_l)$$

$r_j$  – współczynnik korelacji liniowej między zmienną endogeniczną a  $j$ -tą zmienną egzogeniczną,

$r_{ij}$  – współczynnik korelacji liniowej między  $i$ -tą i  $j$ -tą zmienną egzogeniczną występującą w danej kombinacji,

$I_l$  – zbiór numerów zmiennych tworzących  $l$ -tą kombinację.

**Pojemnością integralną nośników informacji**, łączną pojemnością wszystkich zmiennych objaśniających występujących w  $l$ -tej kombinacji, nazywamy wyrażenie:

$$H_l = \sum_{j \in I_l} h_{lj} \quad (l = 1, 2, \dots, C)$$

Pojemność integralna stanowi kryterium doboru kombinacji zmiennych objaśniających do modelu ekonometrycznego. Spośród wszystkich możliwych kombinacji zmiennych wybiera się tą, dla której pojemność integralna jest największa.

## II. KLASYCZNY MODEL REGRESJI

1. Zapis modelu
2. Założenia modelu liniowej regresji wielu zmiennych
3. Klasyczna metoda najmniejszych kwadratów
4. Interpretacja współczynników równania

### 1. ZAPIS MODELU

Rozważamy liniowy model z wieloma zmiennymi objaśniającymi:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i, \quad i = 1, 2, \dots, n$$

gdzie

- $y_i$  –  $i$ -ta obserwacja zmiennej objaśnianej,
- $x_{ji}$  –  $i$ -ta obserwacja  $j$ -tej zmiennej objaśniającej,
- $\varepsilon_i$  – składnik losowy,
- $\beta_0, \dots, \beta_k$  – nieznanne parametry strukturalne modelu,
- $n$  – liczebność próby.

W zapisie wektorowym model ten ma postać:  $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$

gdzie

$$\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{(n \times 1)} \quad \text{– wektor obserwacji zmiennej objaśnianej,}$$

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{21} & \dots & x_{k1} \\ 1 & x_{12} & x_{22} & \dots & x_{k2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{1n} & x_{2n} & \dots & x_{kn} \end{bmatrix}_{(n \times k)} \quad \text{– macierz obserwacji zmiennych objaśniających,}$$

w której kolumna jedynek odpowiada stałej w modelu,

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}_{(k \times 1)} \quad \text{– wektor parametrów strukturalnych,}$$

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}_{(n \times 1)} \quad \text{– wektor składników losowych.}$$

### 2. ZAŁOŻENIA MODELU LINIOWEJ REGRESJI WIELU ZMIENNYCH

Aby estymatory parametrów modelu  $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i$  miały wymagane własności statystyczne, przyjmujemy następujące założenia na temat składnika losowego  $\varepsilon_i$ :



**Założenie 1.** Zmienne objaśniające są nielosowe, ich wartości są ustalonymi liczbami rzeczywistymi.

**Założenie 2.** Wartości oczekiwane składników losowych są równe zeru:  $E(\varepsilon_i) = 0$  dla wszystkich  $i$ .

**Założenie 3.** Składnik losowy jest sferyczny, tzn.:

- wariancje składników losowych są stałe:  $D^2(\varepsilon_i) = \sigma^2$  dla wszystkich  $i$ , czyli składnik losowy jest homoskedastyczny,
- składniki losowe  $\varepsilon_i$  i  $\varepsilon_j$  są od siebie niezależne dla wszystkich  $i \neq j$ , czyli nie występuje autokorelacja składnika losowego.

**Założenie 4.** Składniki losowe mają rozkład normalny, co zapisujemy  $\varepsilon_i : N(0, \sigma^2)$ .

**Założenie 5.** Liczebność próby jest większa niż liczba szacowanych parametrów, czyli  $n > k + 1$ .

**Założenie 6.** Pomędzy wektorami obserwacji zmiennych objaśniających nie istnieje zależność liniowa (jest to założenie o braku współliniowości).

### 3. KLASYCZNA METODA NAJMNIEJSZYCH KWADRATÓW

Poszukujemy estymatora wektora  $\beta$  minimalizującego sumę kwadratów reszt, co jest równoważne minimalizacji następującej formy kwadratowej:

$$Q = \mathbf{e}^T \mathbf{e} = (\mathbf{Y} - \mathbf{Xb})^T (\mathbf{Y} - \mathbf{Xb}) = (\mathbf{Y}^T - \mathbf{b}^T \mathbf{X}^T) (\mathbf{Y} - \mathbf{Xb}) = \\ = \mathbf{Y}^T \mathbf{Y} - \mathbf{Y}^T \mathbf{Xb} - \mathbf{b}^T \mathbf{X}^T \mathbf{Y} + \mathbf{b}^T \mathbf{X}^T \mathbf{Xb} = \mathbf{Y}^T \mathbf{Y} - 2\mathbf{b}^T \mathbf{X}^T \mathbf{Y} + \mathbf{b}^T \mathbf{X}^T \mathbf{Xb}$$

gdzie

$$\mathbf{b} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{bmatrix}_{(k \times 1)} \quad \text{– wektor ocen parametrów strukturalnych,}$$

$$\mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}_{(n \times 1)} \quad \text{– wektor reszt modelowych,}$$

$(\cdot)^T$  oznacza transponowanie macierzy.

Forma  $Q$  osiąga ekstremum w punkcie, w którym jej pochodna się zeruje:  $\frac{\partial Q}{\partial \mathbf{b}} = 0$ .

Stąd:  $-2\mathbf{X}^T \mathbf{Y} + 2\mathbf{X}^T \mathbf{Xb} = 0$ ,

oraz wektor ocen parametrów modelu wyraża wzór:

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y},$$

gdzie  $(\cdot)^{-1}$  oznacza macierz odwrotną.

Przyjmując oznaczenie  $\hat{\mathbf{Y}}$  dla wektora wartości teoretycznych, oszacowany model możemy zapisać jako:

$$\hat{\mathbf{Y}} = \mathbf{Xb}.$$

Powyżej zdefiniowany estymator  $\mathbf{b}$  nosi nazwę **estymatora klasycznej metody najmniejszych kwadratów** – KMNK (ang. ordinary least squares OLS).

**Przykład 1.**

Do opisu produkcji przedsiębiorstwa w mln zł ( $Y$ ) zaproponowano zmienną  $X_1$  – zatrudnienie w osobach,  $X_2$  – nakłady inwestycyjne w mln PLN. Zadaniem jest oszacowanie parametrów modelu liniowego postaci  $y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon_t$  na podstawie danych zawartych w tabeli:

<i>Lata</i>	$y_t$	$x_{1t}$	$x_{2t}$
2001	91	478	78
2002	95	506	84
2003	103	577	95
2004	142	632	94
2005	163	670	104
2006	180	659	119
2007	211	688	144
2008	242	717	202

W celu estymacji modelu wykorzystamy wzór:

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$$

Wyznaczamy macierze  $\mathbf{X}$  i  $\mathbf{Y}$ :

$$\mathbf{X} = \begin{bmatrix} 1 & 478 & 78 \\ 1 & 506 & 84 \\ 1 & 577 & 95 \\ 1 & 632 & 94 \\ 1 & 670 & 104 \\ 1 & 659 & 119 \\ 1 & 688 & 144 \\ 1 & 717 & 202 \end{bmatrix} \quad \mathbf{Y} = \begin{bmatrix} 91 \\ 95 \\ 103 \\ 142 \\ 163 \\ 180 \\ 211 \\ 242 \end{bmatrix}$$

Iloczyn macierzowe  $\mathbf{X}^T \mathbf{X}$  oraz  $\mathbf{X}^T \mathbf{Y}$  są następujące:

$$\mathbf{X}^T \mathbf{X} = \begin{bmatrix} 8 & 4927 & 920 \\ 4927 & 3087487 & 586018 \\ 920 & 586018 & 117518 \end{bmatrix} \quad \mathbf{X}^T \mathbf{Y} = \begin{bmatrix} 1227 \\ 787255 \\ 155851 \end{bmatrix}$$

W wyniku odwrócenia macierzy  $\mathbf{X}^T \mathbf{X}$  otrzymamy:

$$(\mathbf{X}^T \mathbf{X})^{-1} = \begin{bmatrix} 9,906439 & -0,020343 & 0,023887 \\ -0,020343 & 0,000048 & -0,000079 \\ 0,023887 & -0,000079 & 0,000217 \end{bmatrix}$$

Możemy obecnie wyznaczyć wektor ocen parametrów strukturalnych modelu:

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} = \begin{bmatrix} -136,729 \\ 0,3418 \\ 0,692152 \end{bmatrix}$$

Model w postaci oszacowanej można zapisać następująco:



$$\hat{y}_t = -136,729 + 0,3418x_{1t} + 0,692152x_{2t}$$

Dodatkowo wyznaczamy wartości teoretyczne  $\hat{y}_t$  oraz reszty modelowe  $e_t$  i ich kwadraty:

$y_t$	$\hat{y}_t$	$e_t = y_t - \hat{y}_t$	$e_t^2$
91	80,63969	10,3603	107,3361
95	94,363	0,637	0,405765
103	126,2445	-23,2445	540,3061
142	144,3513	-2,3513	5,528818
163	164,2613	-1,2613	1,590796
180	170,8837	9,1163	83,1062
211	198,0997	12,9003	166,4168
242	248,1567	-6,1567	37,90541
Suma	X	0,0001	942,5959

#### 4. INTERPRETACJA WSPÓŁCZYNNIKÓW RÓWNANIA

Oceny estymatorów parametrów modelu ekonometrycznego  $b_j$  ( $j=1,2,\dots,k$ ) informują, jak jednostkowa zmiana  $j$ -tej zmiennej objaśniającej wpłynie na średnią wartość zmiennej objaśnianej, przy założeniu, że pozostałe zmienne pozostaną na tym samym poziomie.

##### Przykład 2.

W modelu  $\hat{y}_t = -136,729 + 0,3418x_{1t} + 0,692152x_{2t}$  z przykładu 1. ocena parametru  $b_2$  ma następującą interpretację:

„Wzrost wartości nakładów inwestycyjnych (zmiennej  $X_2$ ) o 1 mln zł (o jednostkę) wywoła średni przyrost wartości produkcji przedsiębiorstwa (zmiennej  $Y$ ) o 692,152 tys. zł (o 0,692152 jednostki), przy założeniu, że zatrudnienie (zmienna objaśniająca  $X_1$ ) się nie zmieni (założenie *ceteris paribus*).”

### III. WERYFIKACJA MODELU

1. Ocena jakości modelu
2. Badanie istotności zmiennych objaśniających modelu
3. Weryfikacja reszt

#### 1. OCENA JAKOŚCI MODELU

Po oszacowaniu parametrów modelu ekonometrycznego należy przeprowadzić analizę dopasowania modelu do danych empirycznych.

Miary określające stopień zgodności modelu z danymi empirycznymi obliczane są na podstawie reszt  $e_i = y_i - \hat{y}_i$ . Model tym lepiej pasuje do danych empirycznych, im reszty co do wartości bezwzględnej są mniejsze.

O zgodności modelu z danymi empirycznymi mówi wariancja składnika losowego  $\sigma^2$ . Estymatorem wariancji  $\sigma^2$  składnika losowego w liniowym modelu z  $k$  zmiennymi objaśniającymi szacowanym KMNK jest **wariancja reszt**:

$$S_e^2 = \frac{1}{n-k-1} \cdot \sum_{i=1}^n e_i^2$$

**Odchylenie standardowe reszt** wyznaczamy jako:

$$S_e = \sqrt{\frac{1}{n-k-1} \cdot \sum_{i=1}^n e_i^2}$$

**Odchylenie standardowe reszt** (inaczej: błąd estymacji, *standard error of estimation*) mówi, o ile przeciętnie zaobserwowane wartości zmiennej objaśnianej różnią się od wartości teoretycznych tej zmiennej wyznaczonych z modelu.

#### Przykład 3.

Dla danych z przykładu 1:

$$S_e = \sqrt{\frac{1}{n-k-1} \sum_{i=1}^n e_i^2} = \sqrt{\frac{1}{8-2-1} \cdot 942,596} = \sqrt{188,519} = 13,730$$

Wartości empiryczne produkcji przedsiębiorstwa różnią się od wartości teoretycznych przeciętnie o 13,73 mln zł.

Wadą powyższych miar jest fakt, że przeskalowanie zmiennej objaśnianej powoduje zmianę ich wartości.

Dokonyamy teraz **dekompozycji wariancji zmiennej objaśnianej**. Całkowitą zmienność zmiennej objaśnianej można rozbić na dwie części:

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n e_i^2$$

Inaczej zapisując:

$$SST = SSR + SSE$$

$SST = \sum_{i=1}^n (y_i - \bar{y})^2$  to suma kwadratów odchylenia wartości empirycznych  $y$ -ka od wartości średniej, tzw. **całkowita zmienność zmiennej objaśnianej** (*sum of squares-total*),

$SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$  to suma kwadratów odchylenia wartości teoretycznych  $y$ -ka od wartości średniej, tzw. **zmienność zmiennej objaśnianej wyjaśniona przez model** (*sum of squares-*



regression), a  $SSE = \sum_{i=1}^n e_i^2$  - suma kwadratów reszt, czyli **zmienność zmiennej objaśnianej nie wyjaśniona przez model** (*sum of squares-error*).

Dzieląc stronami wzór na całkowitą zmienność  $y$ -ka przez  $SST$  otrzymujemy:  $1 = \frac{SSR}{SST} + \frac{SSE}{SST}$ .  
Stąd wynikają kolejne miary dopasowania modelu do danych empirycznych:

**Współczynnik determinacji  $R^2$  (*R-squared*):**

$$R^2 = \frac{SSR}{SST} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Informuje on, jaka część całkowitej zmienności zmiennej objaśnianej została wyjaśniona przez model.

**Współczynnik zbieżności (indeterminacji)  $\phi^2$ :**

$$\phi^2 = \frac{SSE}{SST} = \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Informuje on, jaka część całkowitej zmienności zmiennej objaśnianej nie została wyjaśniona przez model.

**Współczynnik determinacji** można również wyznaczyć ze wzoru:

$$R^2 = 1 - \phi^2 = 1 - \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Współczynnik  $R^2$  przyjmuje wartości z przedziału  $\langle 0,1 \rangle$ .  $R^2 = 1$  dotyczy doskonałego dopasowania modelu do danych empirycznych. Wtedy wszystkie reszty przyjmują wartość 0. Wyznaczona MNK hiperpłaszczyzna przechodzi wówczas przez wszystkie punkty empiryczne.  $R^2 = 0$  odpowiada modelowi  $\hat{y}_i = b_0$ , w którym żadne zmienne objaśniające nie wyjaśniają zmienności  $y$ -ka. Im większa jest wyjaśniona modelem zmienność zmiennej objaśnianej, tym wartość  $R^2$  bliższa jest jedynce.

W przypadku modelu z jedną zmienną objaśniającą  $R^2$  równy jest kwadratowi współczynnika korelacji z próby między zmienną objaśnianą a objaśniającą.

Dla modelu z wieloma zmiennymi objaśniającymi informacja, jaką niesie wartość  $R^2$  może być fałszywa w wypadku występowania **efektu katalizy**. Wówczas silne skorelowanie zmiennych objaśniających między sobą podwyższa wartość współczynnika determinacji.

Po wprowadzeniu dodatkowej zmiennej objaśniającej do modelu wartość współczynnika determinacji nigdy nie maleje. Trudno więc na podstawie  $R^2$  porównywać modele z taką samą zmienną objaśniającą i różną liczbą zmiennych objaśniających. Zaleca się wtedy stosowanie **skorygowanego współczynnika determinacji** (*adjusted R-squared*):

$$\bar{R}^2 = 1 - \frac{\sum_{i=1}^n e_i^2 / (n-k-1)}{\sum_{i=1}^n (y_i - \bar{y})^2 / (n-1)}$$

Wartości  $\bar{R}^2$  maleją przy wprowadzaniu zmiennych nie wywołujących znaczącego przyrostu wyjaśnienia zmienności zmiennej objaśnianej.

Zależność między  $\bar{R}^2$  a  $R^2$  jest następująca:  $\bar{R}^2 = R^2 - \frac{k}{n-k-1}(1-R^2)$ .

Zachodzi  $\bar{R}^2 \leq R^2$  oraz  $\bar{R}^2 \in (-\infty, 1)$ .

Wśród innych miar dopasowania modelu do danych empirycznych występuje **średni błąd bezwzględny** (*mean absolute error*):

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i|$$

Im mniejsze są wartości  $MAE$ , tym dopasowanie modelu lepsze.

**Współczynnik zmienności losowej**  $W_e = \frac{S_e}{\bar{y}} \cdot 100\%$  określa, jaki procent średniej wartości zmiennej objaśnianej stanowi odchylenie standardowe reszt. Im mniejsza jest wartość  $W_e$ , tym dopasowanie modelu lepsze.

**Współczynnik korelacji wielorakiej**  $R$  określa siłę związku liniowego zmiennej objaśnianej ze wszystkimi zmiennymi objaśniającymi modelu.

$$R = \sqrt{R^2}$$

Ponieważ  $R = r_{y\hat{y}}$ , stąd współczynnik korelacji wielorakiej informuje w jakim stopniu są ze sobą skorelowane empiryczne i teoretyczne wartości zmiennej objaśnianej.

## 2. BADANIE ISTOTNOŚCI ZMIENNYCH OBJAŚNIAJĄCYCH MODELU

W rozdziale niniejszym zajmiemy się wnioskowaniem o parametrach liniowego modelu ekonometrycznego:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i, \quad i = 1, 2, \dots, n$$

Jego oszacowana postać jest następująca:

$$\hat{y}_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + \dots + b_k x_{ki}$$

$b_0, \dots, b_k$  są estymatorami parametrów  $\beta_0, \dots, \beta_k$  mającymi rozkłady normalne.

Wartości oczekiwane tych estymatorów równe są odpowiednio  $\beta_0, \dots, \beta_k$ .

Macierz wariancji i kowariancji estymatora  $\mathbf{b}$  jest równa  $\sigma^2 (\mathbf{X}^T \mathbf{X})^{-1}$ .

Estymatorem macierzy wariancji i kowariancji estymatora  $\mathbf{b}$  jest  $\mathbf{D}^2(\mathbf{b}) = S_e^2 (\mathbf{X}^T \mathbf{X})^{-1}$ .

Niech  $d_{ij}$ ,  $i = 0, 1, \dots, k$ ,  $j = 0, 1, \dots, k$ , oznaczają elementy macierzy  $\mathbf{D}^2(\mathbf{b})$ .

Wtedy elementy diagonalne  $d_{jj}$ ,  $j = 0, 1, \dots, k$ , są ocenami wariancji estymatorów poszczególnych parametrów modelu, natomiast pierwiastki kwadratowe z ocen wariancji są standardowymi błędami szacunku parametrów  $\beta_j$ ,  $j = 0, 1, \dots, k$ :

$$S_{b_j} = \sqrt{d_{jj}}$$



**Standardowy błąd szacunku parametru**  $S_{b_j}$  informuje, o ile średnio ocena  $b_j$  różni się od parametru  $\beta_j$ , jeśli parametr  $\beta_j$  został oszacowany na podstawie różnych prób składających się z tej samej liczby obserwacji.

Standardowy błąd szacunku powinien być jak najmniejszy w stosunku do oceny parametru. W praktyce przyjmuje się, że nie powinien przekraczać 50% jej wartości bezwzględnej, jeśli liczba stopni swobody jest większa niż 20.

**Przykład 4.**

Wyznamy standardowe błędy szacunku parametrów dla modelu  $\hat{y}_t = -136,729 + 0,3418x_{1t} + 0,692152x_{2t}$  z przykładu 1.

$$(\mathbf{X}^T \mathbf{X})^{-1} = \begin{bmatrix} 9,906439 & -0,020343 & 0,023887 \\ -0,020343 & 0,000048 & -0,000079 \\ 0,023887 & -0,000079 & 0,000217 \end{bmatrix}$$

$$S_e = 13,73, \quad S_e^2 = 188,5129$$

$$\mathbf{D}^2(\mathbf{b}) = \begin{bmatrix} 1867,492 & -3,834831 & 4,503038 \\ -3,83483 & 0,009016 & -0,014936 \\ 4,503038 & -0,014936 & 0,040832 \end{bmatrix}$$

$$S_{b_0} = \sqrt{d_{00}} = \sqrt{1867,492} = 43,21448,$$

$$S_{b_1} = \sqrt{d_{11}} = \sqrt{0,009016} = 0,09495,$$

$$S_{b_2} = \sqrt{d_{22}} = \sqrt{0,040832} = 0,202068.$$

$$\hat{y}_t = -136,729 + 0,3418x_{1t} + 0,692152x_{2t}$$

[43,215] [0,095] [0,202]

Np. ocena  $b_0$  różni się od parametru  $\beta_0$  średnio o 43,215.

**Przedział ufności dla parametru strukturalnego**

**Estymacja przedziałowa** polega na wyznaczeniu, na podstawie wyników obserwacji, takiego przedziału liczbowego, o którym, z ustalonym z góry, bliskim jedności, prawdopodobieństwem słuszności sądu, można twierdzić, że zawiera w sobie prawdziwą wartość szacowanego parametru.

Przedział taki nosi nazwę **przedziału ufności**, zaś jego końce, odpowiednio, nazwę **dolnej i górnej granicy ufności**.

Prawdopodobieństwo tego, że przedział ufności pokryje nieznaną wartość szacowanego parametru, nazywa się **współczynnikiem ufności** (ozn.  $1-\alpha$ ). Wartość  $\alpha$  nazywamy **poziomem istotności**. W celu skonstruowania przedziału ufności poziomu istotności przyjmujemy z góry.

Wiemy, że estymator  $b_j$  ma rozkład normalny ze średnią  $\beta_j$  i odchyleniem standardowym  $\sigma_{b_j}$

(co zapisujemy jako  $b_j \sim N(\beta_j, \sigma_{b_j})$ ), zatem  $\frac{b_j - \beta_j}{\sigma_{b_j}} \sim N(0,1)$ .

W praktyce zamiast nieznanego  $\sigma_{b_j}$  stosuje się  $S_{b_j}$ . W związku z tym następująca zmienna ma rozkład t-Studenta o  $(n-k-1)$  stopniach swobody:

$$\frac{b_j - \beta_j}{S_{b_j}} \sim t_{n-k-1}.$$

Aby wyznaczyć przedział ufności dla parametru  $\beta_j, j = 0, 1, \dots, k$ , należy dobrać z tablic rozkładu t-Studenta taką wartość  $t_{\alpha, n-k-1}$ , aby spełniona była relacja:

$$P\left\{\frac{|b_j - \beta_j|}{S_{b_j}} \leq t_{\alpha, n-k-1}\right\} = 1 - \alpha$$

Inaczej:  $P\{b_j - t_{\alpha, n-k-1} \cdot S_{b_j} \leq \beta_j \leq b_j + t_{\alpha, n-k-1} \cdot S_{b_j}\} = 1 - \alpha$

$(1 - \alpha) \cdot 100$ -procentowy przedział ufności dla parametru  $\beta_j$  jest więc postaci:

$$(b_j - t_{\alpha, n-k-1} \cdot S_{b_j}, b_j + t_{\alpha, n-k-1} \cdot S_{b_j}), \quad j = 0, 1, \dots, k.$$

W skrócie wartość krytyczną  $t_{\alpha, n-k-1}$  będziemy oznaczać  $t^*$ .

Długość przedziału ufności zależy od poziomu istotności  $\alpha$ , liczby stopni swobody oraz wielkości standardowych błędów szacunku parametrów. Przedział ufności jest tym węższy im wyższa jest wartość poziomu istotności  $\alpha$ , większa liczba stopni swobody (a więc bardziej liczna próba) oraz niższa wartość standardowego błędu szacunku parametru.

Z praktycznego punktu widzenia będzie nas interesowało, czy skonstruowany przedział ufności zawiera liczbę zero. Będziemy twierdzić, że jeśli do przedziału ufności należy 0, to dany parametr strukturalny modelu jest statystycznie nieistotny.

#### Przykład 5.

$$\hat{y}_t = -136,729 + 0,3418x_{1t} + 0,692152x_{2t}, \quad t = 1, \dots, 8$$

[43,215] [0,095] [0,202]

Wartość krytyczna statystyki t-Studenta dla poziomu istotności  $\alpha=0,05$  i  $n-k-1=8-2-1=5$  stopni swobody wynosi  $t^*=2,5706$ .

Przedział ufności dla parametru  $\beta_1$ :

$$(b_1 - t_{0,05;5} \cdot S_{b_1}; b_1 + t_{0,05;5} \cdot S_{b_1}) = (0,3418 - 2,5706 \cdot 0,095; 0,3418 + 2,5706 \cdot 0,095) =$$

$$= (0,098; 0,586)$$

Można sądzić na 95%, że przedział ten obejmuje nieznaną wartość parametru  $\beta_1$ .

Do przedziału tego nie należy liczba 0, więc można twierdzić, że parametr  $\beta_1$  jest statystycznie istotny.

Zmienna  $X_1$  wywiera tym samym statystycznie istotny wpływ na zmienną  $Y$ .

#### Test t-Studenta na istotność parametru strukturalnego

Zajmiemy się teraz weryfikowaniem hipotezy dotyczącej braku statystycznej istotności parametru  $\beta_j, j = 0, 1, \dots, k$ , stojącego przy zmiennej objaśniającej  $X_j$  w modelu ekonometrycznym. Hipotezy badawcze formułujemy następująco:

$$H_0 : \beta_j = 0 \text{ (parametr jest statystycznie nieistotny)}$$

$$H_1 : \beta_j \neq 0 \text{ (parametr jest statystycznie istotny)}$$

Jeśli hipoteza zerowa jest prawdziwa, to przy założeniu, że składnik losowy modelu ma rozkład normalny, statystyka  $t_j = \frac{b_j}{S_{b_j}}, j = 0, 1, \dots, k$ , ma rozkład t-Studenta z  $n-k-1$  stopniami swobody.

Procedura postępowania:

1. Dla parametru  $\beta_j$  wyznaczamy wartość empiryczną statystyki  $t_j$ .
2. Z tablic dla rozkładu t-Studenta dla zadanego poziomu istotności  $\alpha$  i  $n-k-1$  stopni swobody odczytujemy krytyczną wartość  $t^*$ .
3. Jeżeli  $|t_j| > t^*$ , to hipotezę zerową  $H_0$  odrzucamy na rzecz hipotezy alternatywnej  $H_1$ .  
Mówi się wtedy o **statystycznej istotności parametru**  $\beta_j$ . Uznajemy, że zmienna  $X_j$  wywiera statystycznie istotny wpływ na zmienną  $Y$ .  
Jeżeli  $|t_j| \leq t^*$ , to nie ma podstaw do odrzucenia hipotezy  $H_0$ . Mówi się wtedy o **statystycznej nieistotności parametru**  $\beta_j$ . Uznajemy, że zmienna  $X_j$  nie wywiera statystycznie istotnego wpływu na zmienną  $Y$ .

Jeśli liczba stopni swobody jest większa niż 20, to wtedy dla  $\alpha = 0,05$  wartość krytyczna  $t^* \approx 2$ . W konsekwencji standardowy błąd szacunku parametru nie powinien przekraczać 50% wartości bezwzględnej oceny parametru.

#### Przykład 6.

$$\hat{y}_t = -136,729 + 0,3418x_{1t} + 0,692152x_{2t}, \quad t = 1, \dots, 8$$

[43,215] [0,095] [0,202]

Zweryfikujemy hipotezę  $H_0 : \beta_1 = 0$  wobec hipotezy alternatywnej  $H_1 : \beta_1 \neq 0$ .

Wartość krytyczna statystyki t-Studenta dla  $\alpha = 0,05$  i 5 stopni swobody wynosi  $t^* = 2,5706$ .

Wartość empiryczna statystyki t-Studenta dla  $\beta_1$  wynosi:  $t_1 = \frac{0,3418}{0,095} = 3,598$ .

Ponieważ  $|t_1| > t^*$ , zatem hipotezę zerową odrzucamy na rzecz hipotezy alternatywnej.

Parametr  $\beta_1$  jest statystycznie istotny.

Zmienna  $X_1$  wywiera statystycznie istotny wpływ na zmienną  $Y$ .

**Wartość p** jest poziomem prawdopodobieństwa, przy którym nie ma podstaw do odrzucenia hipotezy zerowej przy obliczonej na podstawie próby wartości empirycznej statystyki  $t_j$ . Przyjmując zazwyczaj poziom  $\alpha = 0,05$ ,  $H_0$  odrzuca się na korzyść  $H_1$ , gdy  $p \leq 0,05$ .

#### Test F

Hipotezy badawcze dotyczą w tym teście zestawu parametrów strukturalnych:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0 \quad (H_0 : R = 0)$$

$$H_1 : \text{co najmniej jeden z } \beta_j \text{ jest różny od zera, } j = 1, 2, \dots, k. \quad (H_1 : R \neq 0)$$

Hipoteza zerowa mówi, że żadna ze zmiennych objaśniających nie wyjaśnia kształtowania się wartości zmiennej objaśnianej.

Przy prawdziwości hipotezy zerowej statystyka  $F = \frac{(n-k-1)}{k} \cdot \frac{R^2}{1-R^2}$  ma rozkład F-

Snedecora z  $m_1 = k$  oraz  $m_2 = n-k-1$  stopniami swobody.

Procedura postępowania:

1. Na podstawie próby obliczamy wartość empiryczną statystyki  $F$ .
2. Dla zadanego poziomu istotności  $\alpha$  oraz dla liczby stopni swobody  $m_1 = k$  i  $m_2 = n-k-1$  odczytujemy z tablic wartość krytyczną  $F^*$ .
3. Jeżeli  $F > F^*$ , to hipotezę  $H_0$  odrzucamy.  
Jeżeli  $F \leq F^*$ , to nie ma podstaw do odrzucenia hipotezy  $H_0$ .

W praktyce za pomocą testu F dokonuje się weryfikacji hipotezy dotyczącej istotności podzbioru parametrów stojących przy zmiennych, o które poszerzono wyjściowy model. Bada się wtedy łączny efekt nowo wprowadzonych zmiennych.

### Przyczyny braku statystycznej istotności parametrów

Brak statystycznej istotności parametru strukturalnego może wynikać z faktycznego braku związku pomiędzy zmienną objaśniającą a zmienną objaśnianą, ale może też być spowodowana innymi przyczynami:

- niską jakością danych statystycznych,
- małą liczebnością próby,
- niewłaściwie dobranym zestawem zmiennych objaśniających,
- niewłaściwą postacią analityczną modelu.

## 3. WERYFIKACJA RESZT

Sprawdzimy teraz, czy w szacowanym modelu ekonometrycznym są spełnione założenia KMNK dotyczące własności składnika losowego. Weryfikację taką przeprowadzimy na podstawie reszt będących oszacowaniami składników losowych.

### Badanie losowości

#### Test liczby serii

Za pomocą testu liczby serii weryfikujemy hipotezę o trafności doboru postaci analitycznej modelu. Przykładowo możemy sprawdzić, czy model  $y_i = f(x_{1i}, \dots, x_{ki}, \varepsilon_i)$ , gdzie  $f$  jest dowolną funkcją (np. liniową), poprawnie opisuje zależność pomiędzy zmienną objaśnianą a zmiennymi objaśniającymi. Stawiamy wówczas hipotezy:

$$H_0: y_i = f(x_{1i}, x_{2i}, \dots, x_{ki}, \varepsilon_i) \text{ (przyjęta postać analityczna modelu jest poprawna)}$$

$$H_1: y_i \neq f(x_{1i}, x_{2i}, \dots, x_{ki}, \varepsilon_i) \text{ (przyjęta postać analityczna modelu nie jest poprawna)}$$

Procedura postępowania:

1. Porządkujemy niemalejąco reszty w próbie według zmiennej porządkującej. Zmienną porządkującą jest zmienna czasowa, gdy dane pochodzą z szeregow czasowych, natomiast dla danych przekrojowych – jedna ze zmiennych objaśniających.
2. Dla uporządkowanego ciągu reszt obliczamy liczbę serii reszt modelu (ozn.  $S$ ). Seria jest sekwencją reszt o takim samym znaku. Jeśli reszty mają charakter losowy, to liczba serii jest zmienną losową.
3. Z tablic testu liczby serii dla liczby reszt dodatnich  $n_1$ , liczby reszt ujemnych  $n_2$  oraz przyjętego poziomu istotności  $\alpha/2$  i  $1-\alpha/2$  odczytujemy krytyczne liczby serii  $S_1^*$  i  $S_2^*$ .
4. Jeśli  $S_1^* < S < S_2^*$ , to nie ma podstaw do odrzucenia hipotezy  $H_0$ , zatem reszty mają charakter losowy. Jeśli  $S \leq S_1^*$  lub  $S \geq S_2^*$  hipotezę  $H_0$  należy odrzucić, rozkład odchyleń uznajemy za nielosowy i twierdzimy, że postać modelu została źle dobrana.

Popularne tablice testu serii skonstruowano dla  $n_1 < 20$  oraz  $n_2 < 20$ . Ze wzrostem liczb  $n_1$  i  $n_2$  rozkład liczby serii dąży do rozkładu normalnego. Wtedy do oceny losowości ciągu reszt wykorzystujemy się statystykę:

$$Z = \frac{S - E(S)}{\sigma_S}, \quad \text{gdzie } E(S) = \frac{2n_1n_2}{n_1 + n_2} + 1, \quad \sigma_S = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}.$$

Statystyka  $Z$  ma asymptotyczny rozkład normalny  $N(0, 1)$ .

### Badanie normalności

#### a) Test Shapiro-Wilka

$H_0$ : składnik losowy modelu  $y = \mathbf{X}\beta + \varepsilon$  ma rozkład normalny

$H_1$ : składnik losowy modelu  $y = \mathbf{X}\beta + \varepsilon$  nie ma rozkładu normalnego



Procedura postępowania:

1. Porządkujemy reszty według wartości niemalejących, otrzymując ciąg  $e_{(1)}, e_{(2)}, \dots, e_{(n)}$ .

$$2. \text{ Obliczamy wartość statystyki } W = \frac{[\sum_{i=1}^{\lfloor n/2 \rfloor} a_{n,i} \cdot (e_{(n-i+1)} - e_{(i)})]^2}{\sum_{i=1}^n e_i^2},$$

gdzie  $\lfloor n/2 \rfloor$  – część całkowita liczby  $n/2$ ,

$a_{n,i}$  – współczynnik Shapiro-Wilka odczytany z tablic.

3. Z tablic testu Shapiro-Wilka dla poziomu istotności  $\alpha$  i wielkości próby  $n$  odczytujemy wartość krytyczną  $W^*$ .

4. Jeżeli  $W \geq W^*$ , to nie ma podstaw do odrzucenia hipotezy  $H_0$ .  
Jeżeli  $W < W^*$ , hipotezę  $H_0$  odrzucamy na korzyść hipotezy  $H_1$ .

Test ten może być stosowany dla małych prób (np. gdy  $n \leq 30$ ).

### b) Test Jarque'a-Bery

Test ten polega na porównaniu, jak wartości miar asymetrii i kurtozy, obliczonych na podstawie reszt, różnią się od wartości tych miar dla rozkładu normalnego.

$H_0$ : składnik losowy modelu  $y = \mathbf{X}\beta + \varepsilon$  ma rozkład normalny

$H_1$ : składnik losowy modelu  $y = \mathbf{X}\beta + \varepsilon$  nie ma rozkładu normalnego

Procedura postępowania:

1. Na podstawie reszt modelowych obliczamy:

**współczynnik asymetrii:**

$$A = \frac{M_3}{S^3} = \frac{\frac{1}{n} \sum_{i=1}^n e_i^3}{\left( \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2} \right)^3}$$

$A > 0$  w przypadku asymetrii prawostronnej.

$A < 0$  w przypadku asymetrii lewostronnej.

$A = 0$  dla rozkładów symetrycznych (m.in. dla rozkładu normalnego).

**współczynnik skupienia (kurtozę):**

$$K = \frac{M_4}{S^4} = \frac{\frac{1}{n} \sum_{i=1}^n e_i^4}{\left( \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2} \right)^4}.$$

$K < 3$ , jeśli rozkład jest bardziej spłaszczony niż rozkład normalny.

$K > 3$ , jeśli rozkład jest bardziej wysmukły niż rozkład normalny.

Dla rozkładu normalnego  $K = 3$ .

2. Wyznaczamy wartość statystyki  $JB = n \left( \frac{1}{6} A^2 + \frac{1}{24} (K - 3)^2 \right)$ . Ma ona asymptotyczny rozkład  $\chi^2$  z 2 stopniami swobody.

3. Z tablic dla rozkładu  $\chi^2$  dla poziomu istotności  $\alpha$  i 2 stopni swobody odczytujemy wartość krytyczną  $\chi^2_*$ .

4. Jeśli  $JB > \chi^2_*$ , to hipotezę  $H_0$  o normalności składnika losowego odrzucamy.

Jeśli  $JB \leq \chi^2_*$ , nie ma podstaw do odrzucenia  $H_0$ .

Test Jarque'a-Bery stosujemy jedynie dla dużych prób.

Odrzucenie hipotezy o normalności składnika losowego oznacza, że estymatory parametrów modelu nie mają rozkładu normalnego. Należy wtedy ostrożnie wnioskować na podstawie testów, wykorzystujących założenie normalności składnika losowego (np. testu t-Studenta).

### Badanie autokorelacji

Negatywne zjawisko autokorelacji składników losowych może się pojawić na skutek błędnej specyfikacji modelu (brak ważnej zmiennej objaśniającej, niewłaściwa postać analityczna) lub z powodu powolnego wygasania efektów czynników przypadkowych zaburzających normalny przebieg prawidłowości ekonomicznych.

Mówimy, że **nie występuje autokorelacja składników losowych** (są one nieskorelowane), jeśli  $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$  dla  $i \neq j, i, j = 1, 2, \dots, n$ .

Jeśli  $\text{cov}(\varepsilon_i, \varepsilon_j) \neq 0$  dla  $j=i-p, i=p+1, \dots, n$ , to mówimy, że **wystąpiła autokorelacja rzędu  $p$** .

Jeśli  $\text{cov}(\varepsilon_i, \varepsilon_{i-1}) \neq 0$  dla  $i=2, \dots, n$ , to **wystąpiła autokorelacja I-ego rzędu**.

Miernikiem autokorelacji składników losowych rzędu  $p$  są współczynniki autokorelacji z próby, będące współczynnikami korelacji liniowej pomiędzy resztami odległymi od siebie o  $p$  okresów.

Do badania autokorelacji I-ego rzędu wykorzystuje się następujący współczynnik  $\hat{\rho}$ :

$$\hat{\rho} = \frac{\sum_{i=2}^n e_i e_{i-1}}{\sqrt{\sum_{i=1}^n e_i^2 \cdot \sum_{i=2}^n e_{i-1}^2}}$$

Współczynnik  $\hat{\rho}$  przyjmuje wartości z przedziału  $\langle -1, 1 \rangle$ , jego znak określa kierunek zależności, a bezwzględna wartość wskazuje na siłę zależności.

**a) Test Durбина-Watsona** weryfikuje istnienie autokorelacji pierwszego rzędu.

W teście tym stawiamy hipotezy następująco:

$H_0: \rho = 0$  (nie występuje autokorelacja I-ego rzędu składników losowych)

$H_1: \rho > 0$  (ma miejsce dodatnia autokorelacja I-ego rzędu składników losowych)

albo

$H_0: \rho = 0$  (nie występuje autokorelacja I-ego rzędu składników losowych)

$H_1: \rho < 0$  (ma miejsce ujemna autokorelacja I-ego rzędu składników losowych)

Sprawdzianem jest statystyka Durбина-Watsona postaci:

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \approx 2(1 - \hat{\rho})$$

Statystyka  $d$  przyjmuje wartości z przedziału  $\langle 0, 4 \rangle$ .

Wartość  $d = 2$  świadczy o braku autokorelacji (wtedy  $\hat{\rho} = 0$ ).

Gdy  $d \approx 0$ , to ma miejsce silna autokorelacja dodatnia (wtedy  $\hat{\rho} \approx 1$ ).

Gdy  $d \approx 4$ , to ma miejsce silna autokorelacja ujemna (wtedy  $\hat{\rho} \approx -1$ ).

Rozkład statystyki  $d$  zależy od liczby obserwacji i liczby szacowanych parametrów. Tablice zawierają wartości krytyczne: górną  $d_U$  i dolną  $d_L$ .



Wnioskowanie na podstawie  $d$ :

1. Jeśli  $\hat{\rho} > 0$  (czyli  $d < 2$ ), to weryfikuje się hipotezy:

$$H_0: \rho = 0 \quad \text{wobec} \quad H_1: \rho > 0$$

Gdy  $d > d_U$ , to brak podstaw do odrzucenia hipotezy  $H_0$ .

Gdy  $d < d_L$ , to odrzucamy hipotezę  $H_0$  i ma miejsce autokorelacja dodatnia.

Gdy  $d_L \leq d \leq d_U$  brak rozstrzygnięcia testu.

2. Jeśli  $\hat{\rho} < 0$  (czyli  $d > 2$ ), to weryfikuje się hipotezy:

$$H_0: \rho = 0 \quad \text{wobec} \quad H_1: \rho < 0$$

Obliczamy  $d' = 4 - d$ .

Gdy  $d' > d_U$ , to brak podstaw do odrzucenia hipotezy  $H_0$ .

Gdy  $d' < d_L$ , to odrzucamy hipotezę  $H_0$  i ma miejsce autokorelacja ujemna.

Gdy  $d_L \leq d' \leq d_U$  brak rozstrzygnięcia testu.

Testu Durбина-Watsona nie stosuje się, jeśli w modelu jako zmienna objaśniająca występuje opóźniona zmienna objaśniana.

### b) Test mnożnika Lagrange'a

$H_0: \rho = 0$  (nie występuje autokorelacja I-ego rzędu składników losowych)

$H_1: \rho \neq 0$  (ma miejsce dodatnia autokorelacja I-ego rzędu składników losowych)

Procedura postępowania:

1. Wyznaczamy reszty  $e_i$  modelu  $\hat{y}_i = b_0 + b_1x_{1i} + b_2x_{2i} + \dots + b_kx_{ki}$ ,  $i = 1, 2, \dots, n$ .

2. Szacujemy model pomocniczy  $\hat{e}_i = a_0 + a_1x_{1i} + \dots + a_kx_{ki} + a_{k+1}e_{i-1}$ ,  $i = 2, 3, \dots, n$  i wyznaczamy współczynnik determinacji  $R^2$  dla tego modelu.

3. Obliczamy  $(n-1) \cdot R^2$  i porównujemy ją z wartością krytyczną testu  $\chi^2$  dla poziomu istotności  $\alpha$  i 1 stopnia swobody  $\chi^2_*$ .

4. Jeśli  $(n-1) \cdot R^2 > \chi^2_*$ , to odrzucamy  $H_0$  (wystąpiła autokorelacja pierwszego rzędu).

Jeśli  $(n-1) \cdot R^2 \leq \chi^2_*$ , to nie ma podstaw do odrzucenia  $H_0$  o braku autokorelacji.

Test mnożnika Lagrange'a może być stosowany jedynie dla dużej liczby obserwacji.

Gdy mamy do czynienia z autokorelacją składnika losowego, estymatory wariancji parametrów strukturalnych szacowanych KMNK są obciążone. W rezultacie fałszywych informacji dostarczają wartości statystyk t-Studenta oraz przedziały ufności. W takiej sytuacji parametry modelu należy szacować uogólnioną metodą najmniejszych kwadratów lub metodą różniczki zupełnej.

### Badanie homoskedastyczności

W KMNK zakłada się, że wariancje składników losowych są stałe:  $D^2(\varepsilon_i) = \sigma^2$  dla  $i = 1, 2, \dots, n$  (składnik losowy jest **homoskedastyczny**). W praktyce, zwłaszcza w modelach dla danych przekrojowych, założenie to często nie jest spełnione - składnik losowy jest **heteroskedastyczny** (np. model wydatków gospodarstw w zależności od ich dochodów).

### Test Goldfelda-Quandt

Dla dwóch części populacji, co do których przypuszczamy, że cechują je różne wariancje składnika losowego, formułujemy następujące hipotezy:

$$H_0: \sigma_1^2 = \sigma_2^2 \quad (\text{składnik losowy jest homoskedastyczny})$$

$$H_1: \sigma_1^2 \neq \sigma_2^2 \quad (\text{składnik losowy jest heteroskedastyczny})$$

gdzie  $\sigma_1^2, \sigma_2^2$  - wariancje składników losowych w pierwszej i drugiej części.

Procedura postępowania:

1. Porządkujemy niemalejąco obserwacje w próbie według zmiennej porządkującej. Zmienną porządkującą może być zmienna czasowa lub – dla danych przekrojowych – zmienna objaśniająca podejrzana o wywoływanie heteroskedastyczności.
2. Wybieramy dwie skrajne podpróby. Pominięta liczba obserwacji nie powinna przekraczać  $1/3$  liczebności całej próby. Niech  $n_1$  - liczba obserwacji w pierwszej podpróbie, a  $n_2$  - liczba obserwacji w drugiej. Szacujemy model indywidualnie w każdej podpróbie i wyznaczamy odpowiednie wariancje resztowe  $S_1^2$  i  $S_2^2$ .
3. Obliczamy  $F = S_2^2 / S_1^2$  (w liczniku większa z wariancji). Jeśli hipoteza  $H_0$  jest prawdziwa, to  $F$  ma rozkład F-Snedecora.
4. W tablicach statystycznych dla poziomu istotności  $\alpha$  oraz  $m_1 = n_2 - k - 1$ ,  $m_2 = n_1 - k - 1$  stopni swobody odczytujemy wartość krytyczną  $F^*$ .
5. Jeśli  $F \leq F^*$ , to brak podstaw do odrzucenia hipotezy  $H_0$ .  
Jeśli  $F > F^*$ , to hipotezę  $H_0$  odrzucamy na rzecz hipotezy  $H_1$ .

Jeśli heteroskedastyczność została spowodowana łącznie przez kilka zmiennych objaśniających, stosujemy inne testy, np. **test White'a** lub **Harveya-Godfrey'a**.

W wypadku heteroskedastyczności składnika losowego estymatory parametrów strukturalnych uzyskane KMNK mają obciążoną wariancję, co uniemożliwia poprawną weryfikację hipotez dotyczących wartości parametrów strukturalnych. W takiej sytuacji do estymacji parametrów modelu należy zastosować ważoną MNK.



## IV. ANALIZA ROZWOJU ZJAWISK W CZASIE

1. Szeregi czasowe
2. Modele trendu
3. Wygładzanie szeregu czasowego metodą średniej ruchomej
4. Stacjonarność procesów stochastycznych
5. Modele AR, MA, ARMA i ARIMA
6. Inne modele ze zmiennymi opóźnionymi
7. Kointegracja szeregów czasowych

### 1. SZEREGI CZASOWE

**Szeregiem czasowym** nazywa się uporządkowany według czasu zbiór obserwacji statystycznych.

W wypadku szeregu czasowego przyjmuje się, że wartości zmiennej charakteryzującej określone zjawisko (zmiennej zależnej) zależą od czasu (zmiennej niezależnej).

Rozróżniamy dwa typy szeregów czasowych:

- **szeregi okresów** - zawierają dane dotyczące kształtowania się poziomu zjawiska w całym okresie przyjętym za jednostkę czasu. Istnieje możliwość wydłużania okresów, w jakich dokonuje się pomiaru, np. przechodzenie z danych miesięcznych na kwartalne i roczne, za pomocą agregacji danych.
- **szeregi momentów** - zawierają informacje jedynie o poziomie zjawiska w wyróżnionych momentach czasu (tzn. w chwili dokonywania pomiaru) i nie wiemy, jak się ono kształtowało między kolejnymi momentami obserwacji. Można jednak wyznaczyć przeciętny poziom zjawiska na podstawie pomiarów przeprowadzonych w kilku momentach należących do okresu przyjętego za nową jednostkę czasu (np. mając dane o liczbie zatrudnionych w kolejnych miesiącach można podać średni poziom zatrudnienia w ciągu roku).

**Celem analizy szeregów czasowych** jest:

- oszacowanie parametrów wybranego modelu kształtowania się zmiennej i ocena dokładności dopasowania modelu do danych empirycznych,
- wykorzystanie oszacowanego modelu dla prognozy kształtowania się zjawiska w przyszłych okresach.

Z punktu widzenia działania różnego rodzaju przyczyn, które wywołują określone zmiany w rozwoju zjawiska, wyróżniamy:

- **przyczyny działające w sposób trwały**, powodując wystąpienie określonej tendencji rozwojowej (trendu),
- **przyczyny zakłócające** tą tendencję w sposób **regularny lub nieregularny**.

**Dekompozycja szeregu czasowego** polega na stwierdzeniu występowania w szeregu czasowym takich składników, jak:

- stały (przeciętny) poziom,
- tendencja rozwojowa (trend),
- wahania okresowe,
- wahania cykliczne,
- wahania przypadkowe.

**Tendencja rozwojowa** to systematyczne, jednokierunkowe zmiany (wzrost lub spadek) poziomu badanego zjawiska zachodzące w długim okresie. Wyrównaną linię wyznaczającą zasadniczy kierunek rozwojowy badanego procesu nazywamy **trendem**.

**Wahania okresowe (periodyczne)** są to rytmiczne wahania o określonym cyklu (okresie przebiegu). Najczęściej obserwuje się wahania o cyklu rocznym, przy czym podokresami cyklu w takim przypadku mogą być półrocza, kwartały, miesiące, a nawet dni.

**Wahania sezonowe** – wahania o okresie rocznym. U podstaw występowania wahań sezonowych znajdują się takie czynniki egzogeniczne, jak klimatyczno-przyrodnicze lub kalendarzowe (np. czas trwania dnia lub nocy, temperatura, opady).

**Wahania cykliczne** to długookresowe, rytmiczne wahania wartości zmiennej wokół stałego (przeciętnego) poziomu lub wokół trendu badanej zmiennej.

**Wahania koniunkturalne** to systemowe, falowe wahania rozwoju gospodarki obserwowane w dłuższych od roku okresach. Analiza tego rodzaju wahań wymaga wieloletnich obserwacji.

**Wahania przypadkowe** – powodują pewne odchylenia od zmian regularnych. Są wynikiem działania przyczyn ubocznych, ujawniających się nieregularnie.

W celu identyfikacji poszczególnych składowych szeregu czasowego dla konkretnej zmiennej wykorzystuje się **wykresy** zaobserwowanych danych. Analiza wykresu szeregu czasowego pozwala również na wykrycie **obserwacji nietypowych** oraz **punktów zwrotnych**.

Prawidłowości rozwoju zjawiska wykryte w szeregach czasowych można wyrazić modelem opisującym zależność poziomu zjawiska od czasu, czyli w postaci funkcji:

$$y_t = f(t) + g_l(t) + \varepsilon_t \quad \text{lub} \quad y_t = f(t) \cdot g_l(t) \cdot 10^{\varepsilon_t},$$

gdzie:

$y_t$  – poziom badanego zjawiska zaobserwowany w chwili  $t$ ,

$f(t)$  – funkcja trendu,

$g_l(t)$  – funkcja wahań okresowych dla  $l=1,2,\dots,L$  gdzie  $L$  oznacza liczbę podokresów składowej periodycznej,

$\varepsilon_t$  – zmienna losowa, charakteryzująca efekty oddziaływania wahań przypadkowych.

Modele powyższe nazywa się modelami wahań w czasie odpowiednio typu **addytywnego** i **multiplikatywnego**.

**Dekompozycję szeregu czasowego** na poszczególne składowe przeprowadza się stosując odpowiednie **metody**. Wyodrębnianie tendencji rozwojowej można przeprowadzić za pomocą metody:

1. analitycznej, polegającej na oszacowaniu funkcji trendu,
2. mechanicznej, opartej na średnich.

## 2. MODELE TRENDU

**Modele trendu**, zwane również **modelami tendencji rozwojowej**, zawierają tylko jedną zmienną objaśniającą, którą jest czas  $t$ , czyli:

$$y_t = f(t) + \varepsilon_t,$$

gdzie:

$f$  - symbol dowolnej funkcji,

$t$  - zmienna czasowa przyjmująca najczęściej wartości  $t=1,2,\dots,T$ ,



$y_t$  - zmienna objaśniana,

$\varepsilon_t$  - składnik losowy.

Zadanie wyznaczenia funkcji  $f(t)$  jest nazywane **wygładzaniem** (wyrównywaniem) szeregu czasowego.

Analityczna metoda wyrównywania szeregów czasowych polega na takim dopasowaniu funkcji  $f(t)$  do danych empirycznych, aby jak najlepiej obrazowała ona ogólną tendencję rozwojową, eliminując wahania okresowe i przypadkowe.

W zależności od postaci analitycznej funkcji  $f$  wyróżnia się różne rodzaje trendu. Najczęściej w praktyce stosuje się:

- **funkcję liniową**  $f(t) = \beta_0 + \beta_1 t$

Postać trendu liniowego stosowana jest w przypadku, gdy można przyjąć założenie o stałych przyrostach wartości zmiennej  $y$  w jednostce czasu. Wtedy parametr  $\beta_1$  wyraża stały przyrost z okresu na okres wartości zmiennej objaśnianej. Funkcję powyższą można oszacować za pomocą MNK.

- **funkcję wielomianową**  $f(t) = \beta_0 + \beta_1 t + \dots + \beta_n t^n$

Każdą funkcję ciągłą i ograniczoną można w skończonym przedziale aproksymować wielomianem odpowiedniego stopnia. Korzystając więc z funkcji wielomianowej trendu możemy uzyskać wysoką zgodność modelu z obserwacjami empirycznymi. Nie ma jednak gwarancji, że otrzymana funkcja trendu równie dobrze opisze zmiany zmiennej  $y$  w przyszłości. W wypadku wielomianowej funkcji trendu stosując odpowiednie podstawienia sprowadzamy model do postaci liniowej.

- **funkcję potęgową**  $f(t) = \beta_0 t^{\beta_1}$

Tę postać funkcji trendu stosuje się najczęściej w przypadku, gdy wykładnik  $\beta_1$  jest większy od zera, lecz mniejszy od jedności, wtedy funkcja powyższa charakteryzuje się malejącym tempem wzrostu. Jest to przykład funkcji nieliniowej względem parametrów. Aby sprowadzić ją do postaci liniowej, logarytmujemy obie strony wyrażenia, a następnie szacujemy parametry MNK.

- **funkcję wykładniczą**  $f(t) = e^{\beta_0 + \beta_1 t}$

Własnością charakterystyczną tej funkcji trendu są stałe w czasie przyrosty względne. Przed estymacją funkcję sprowadzamy ją do postaci liniowej za pomocą logarytmowania obu stron.

- **funkcję logarytmiczną**  $f(t) = \beta_0 + \beta_1 \ln t$

Trend logarytmiczny wybieramy, gdy wzrost badanej zmiennej jest coraz wolniejszy. Po wprowadzeniu podstawienia  $t' = \ln t$  otrzymujemy równanie liniowe, którego parametry szacuje się MNK. Jest to więc przykład trendu nieliniowego, ale liniowego względem parametrów.

- **funkcję logistyczną**  $f(t) = \frac{\beta_1}{1 + e^{\beta_2 - \beta_3 t}}$

### Klasyczne podejście do określenia stopnia wielomianu będącego funkcją trendu.

Rozpatrzmy ciąg wielomianów postaci:

$$f_0(t) = \beta_{00}$$

$$f_1(t) = \beta_{10} + \beta_{11} t$$

$$f_2(t) = \beta_{20} + \beta_{21} t + \beta_{22} t^2$$

.....

Każdemu z tych wielomianów odpowiada model trendu postaci:

$$y_t = f_i(t) + \varepsilon_t \quad i \in N$$

Każdy z tych modeli może być szacowany MNK.

Dla modelu trendu wielomianowego stopnia  $i$  wyznacza się wariancję składnika resztowego:

$$S_i^2 = \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n - i - 1}$$

Formułujemy hipotezę  $H_0$  orzekającą, że wariancja  $\sigma_{p-1}^2$  składnika losowego modelu odpowiadającego wielomianowi  $f_{p-1}(t)$  różni się od wariancji  $\sigma_p^2$  składnika losowego modelu odpowiadającego wielomianowi  $f_p(t)$  jedynie wskutek przyczyn losowych, a zatem:

$$H_0 : \sigma_{p-1}^2 = \sigma_p^2$$

Jeżeli hipoteza  $H_0$  jest prawdziwa, to zmienna losowa  $F = S_{p-1}^2 / S_p^2$  ma rozkład F-Snedecora, przy czym  $S_i^2$  oznacza oszacowanie wariancji składnika losowego  $\varepsilon_t$  w modelu trendu wielomianowego stopnia  $i$ .

Obieramy poziom istotności  $\alpha$  i z tablic odczytujemy wartość krytyczną  $F^*$ .

Jeżeli  $F > F^*$ , to uznajemy, że stopień wielomianu nie jest jeszcze wyznaczony i kontynuujemy badanie (posługujemy się ilorazem  $S_p^2$  przez  $S_{p+1}^2$ ).

Jeżeli  $F \leq F^*$ , to uznajemy, że szukanym wielomianem jest  $f_{p-1}(t)$ .

### 3. WYGLĄDZANIE SZEREGU CZASOWEGO METODĄ ŚREDNIEJ RUCHOMEJ

Idea wyrównywania szeregu czasowego za pomocą średnich ruchomych polega na zastąpieniu pierwotnych wartości zmiennej średnimi arytmetycznymi, obliczonymi sekwencyjnie dla wybranej liczby obserwacji.

Najłatwiej wyznacza się tzw. **prostą średnią ruchomą**:

$$\bar{y}_{t-\frac{k-1}{2}} = \frac{1}{k} \sum_{i=t-k+1}^t y_i \quad \text{dla } t \geq k$$

Wyznaczone wartości średnie przyporządkowuje się środkowym obserwacjom, na których podstawie obliczono średnie. Średnie te bywają nazywane średnimi łańcuchowymi ze względu na wzajemne zazębianie się okresów biorących udział w liczeniu.

3-okresowa średnia ruchoma dla  $T$ -elementowego szeregu czasowego:

$$\bar{y}_2 = \frac{y_1 + y_2 + y_3}{3}, \quad \bar{y}_3 = \frac{y_2 + y_3 + y_4}{3}, \quad \dots, \quad \bar{y}_{T-1} = \frac{y_{T-2} + y_{T-1} + y_T}{3}$$

W sytuacji, kiedy średnie ruchome mają być wykorzystane do eliminacji wahań okresowych, muszą obejmować ściśle określona liczbę jednostek czasu, odpowiadającą zaobserwowanemu cyklowi wahań. Na przykład pojawienie się wahań przy danych kwartalnych oznacza potrzebę zastosowania 4-okresowej średniej ruchomej, a wahań dla danych miesięcznych – średniej 12-okresowej.

Pojawia się wtedy problem liczenia średnich ruchomych z parzystej liczby okresów, są to tzw. **średnie scentrowane**, które wyznacza się zgodnie z relacją:

$$\bar{y}_{t-\frac{k}{2}} = \frac{1}{k} \left( \frac{1}{2} \cdot y_{t-k} + \sum_{i=t-k+1}^{t-1} y_i + \frac{1}{2} y_t \right) \quad \text{dla } t \geq k+1$$

gdzie:

$\bar{y}_i$  – wartości średnich ruchomych obliczone dla kolejnych podokresów,

$k$  – krok uśredniania.



Przykładowo średnie scentrowane 4- okresowe wyznaczmy następująco:

$$\bar{y}_3 = \frac{0,5 \cdot y_1 + y_2 + y_3 + y_4 + 0,5 \cdot y_5}{4}, \dots, \bar{y}_{T-2} = \frac{0,5 \cdot y_{T-4} + y_{T-3} + y_{T-2} + y_{T-1} + 0,5 \cdot y_T}{4}$$

Wady mechanicznej metody wyrównywania szeregów czasowych:

- skracanie szeregów czasowych,
- brak możliwości przedstawienia trendu w formie matematycznej funkcji trendu.

#### 4. STACJONARNOŚĆ PROCESÓW STOCHASTYCZNYCH

**Proces stochastyczny**  $\{Y_t\}$  jest rodziną zmiennych losowych o wartościach rzeczywistych indeksowanych przez  $t$ , gdzie  $t$  oznacza czas.

Każdy element  $Y_1, Y_2, \dots, Y_t$  procesu stochastycznego  $\{Y_t\}$  jest zmienną losową.

**Szereg czasowy** jest jedną z realizacji procesu stochastycznego.

Wśród modeli szeregów czasowych najczęściej stosowane są modele opisujące **procesy stacjonarne**. Procesy stacjonarne charakteryzują się tym, że mają stałą wariancję, a ich wartości w poszczególnych momentach czasu oscylują wokół pewnego, względnie stałego poziomu, który jest poziomem średnim dla całego badanego okresu.

Proces stochastyczny  $\{Y_t\}$  jest **silnie stacjonarny**, jeżeli dla dowolnych wartości  $t_1, t_2, \dots, t_n$  i dla dowolnego  $k$ , łączny rozkład prawdopodobieństwa dla  $\{Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}\}$  jest identyczny z łącznym rozkładem prawdopodobieństwa dla  $\{Y_{t_1+k}, Y_{t_2+k}, \dots, Y_{t_n+k}\}$ .

Proces stochastyczny  $\{Y_t\}$  jest **słabo stacjonarny**, jeśli jego średnia oraz wariancja są stałe i niezależne od czasu, natomiast kowariancje pomiędzy równoległymi elementami szeregu zależą jedynie od odległości tych elementów.

$$\begin{aligned} E(Y_t) &= \mu \\ \text{Var}(Y_t) &= \sigma^2 \\ \text{Cov}(Y_t, Y_{t+h}) &= \gamma(h) \end{aligned}$$

Badana zmienna  $Y_t$  jest **stacjonarna**, gdy w modelu:

$$Y_t = \alpha + \beta Y_{t-1} + \varepsilon_t \quad (*)$$

parametr  $\beta$  spełnia nierówność:  $|\beta| < 1$ .

Jeżeli  $|\beta| \geq 1$ , to zmienna  $Y_t$  nie jest stacjonarna.

Jeżeli  $|\beta| > 1$ , to mamy do czynienia ze zjawiskiem eksplodującym w czasie.

Jeżeli  $\beta = 1$  oraz  $\alpha = 0$ , wówczas mamy do czynienia z **procesem błędzenia przypadkowego** (*random walk without drift*).

Jeżeli  $\beta = 1$  oraz  $\alpha \neq 0$ , wówczas mamy do czynienia z **procesem błędzenia przypadkowego z dryfem** (*random walk with drift*).

Na potrzeby testowania stacjonarności zmiennej  $Y_t$  w praktyce korzystniej jest równanie (\*) przekształcić do postaci:

$$\Delta Y_t = \alpha + (\beta - 1)Y_{t-1} + \varepsilon_t \quad (**)$$

Stacjonarność  $Y_t$  zachodzi, gdy  $(\beta - 1) \in (-2; 0)$ . Brak stacjonarności z pierwiastkiem jednostkowym (*unit root nonstationarity*) ma miejsce w pozostałych przypadkach.

### Test Dickeya-Fullera (DF) dla pierwiastka jednostkowego

Przy weryfikacji stacjonarności dla modelu typu (\*) wykorzystywany jest **test DF**.

$H_0 : (\beta - 1) = 0$  proces niestacjonarny; zintegrowany w stopniu co najmniej 1 (I(1))

$H_1 : (\beta - 1) < 0$  proces stacjonarny; I(0)

Szacujemy model (\*\*). Weryfikacja stacjonarności odbywa się przez określenie statystycznej istotności parametru  $(\beta - 1)$ . Wyznaczamy więc wartość statystyki testowej:

$$DF_{emp} = \frac{b-1}{S_{b-1}}.$$

W tablicach dla testu DF znajdujemy dwie wartości krytyczne:  $DF_d$  oraz  $DF_g$  (dla ilości obserwacji  $n$ , poziomu istotności  $\alpha$  i liczby opóźnień  $m=0$ ).

Jeżeli  $DF_{emp} < DF_d$ , to  $H_0$  odrzucamy.

Jeżeli  $DF_{emp} > DF_g$ , to brak podstaw do odrzucenia  $H_0$ .

Jeżeli  $DF_d \leq DF_{emp} \leq DF_g$ , to brak decyzji.

Wadą testu DF jest jego słaba moc. Oznacza to, że możemy wskazać niestacjonarność w przypadku, gdy ona nie występuje (tzw. prawdopodobieństwo błędu II rodzaju). Ewentualny brak stacjonarności w modelach ekonomicznych jest niekorzystny. Dlatego zmienne niestacjonarne, czyli z pierwiastkiem jednostkowym, powinny być wyłączone z szacowanego modelu.

Jeżeli w teście DF brak podstaw do odrzucenia  $H_0$ , w dalszej kolejności w analogiczny sposób weryfikujemy stacjonarność zmiennej  $\Delta Y_t = Y_t - Y_{t-1}$ .

Tym razem badamy model postaci  $\Delta \Delta Y_t = \alpha + (\beta - 1)\Delta Y_{t-1} + \varepsilon_t$ .

Brak stacjonarności dla pewnej zmiennej nie wyklucza stacjonarności dla jej dalszych przyrostów.

W praktyce powyżej przedstawiona weryfikacja jest niewystarczająca do wskazania występowania pierwiastka jednostkowego. Niestacjonarny szereg  $Y_t$  może mieć różne źródła braku stacjonarności; na przykład liniowy **trend deterministyczny**, jak i występowanie pierwiastka jednostkowego (**trend stochastyczny**).

Dla modelu (\*) postać uwzględniająca trend jest następująca:  $Y_t = \alpha + \beta Y_{t-1} + \delta t + \varepsilon_t$ .

Weryfikacja modelu z trendem ma wiele analogii do przypadku wcześniej omawianego modelu bez trendu. W przypadku niestacjonarnej zmiennej  $Y_t$ , która okaże się stacjonarnym procesem w modelu uwzględniającym trend, możemy powiedzieć, iż taki proces jest **stacjonarny względem trendu** (*trend stationary*).

### Test ADF (Augmented Dickey-Fuller)

W przypadku jeżeli pewna zmienna ( $Y_t$ ) występuje w modelu z ilością opóźnień równą  $p$ , to jej stacjonarność powinna być weryfikowana w oparciu o model autoregresyjny.

Dla modelu, który uwzględnia trend oraz zmienną  $Y_t$  z seryjną korelacją, podstawą do weryfikacji jest następująca formuła:

$$\Delta Y_t = \alpha + (\beta - 1)Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \dots + \gamma_p \Delta Y_{t-p} + \delta t + \varepsilon_t$$

Znajdujemy optymalne opóźnienie  $p$  dla zmiennej  $Y$  oraz określamy istotność stałej i trendu. Następnie uwagę poświęcamy testowaniu parametru  $(\beta - 1)$ .



## Stopień integracji

Jeżeli pewna zmienna jest stacjonarna, to jest zintegrowana w stopniu zero (zapis:  $I(0)$ ). Każdy proces ze względu na poziom integracji możemy zapisać, jako  $I(d)$ , gdzie  $d$  oznacza stopień integracji. Stopień integracji oznacza ilość koniecznych do wyznaczenia różnic dla danej zmiennej, aby osiągnąć  $I(0)$ .

Przykładowo, jeśli  $Y_t$  - zmienna niestacjonarna, a  $\Delta Y_t$  - zmienna stacjonarna, to:  $\Delta Y_t$  jest szeregiem zintegrowanym w stopniu  $d = 0$ ;  $\Delta Y_t \sim I(0)$ ,  $Y_t$  jest procesem zintegrowanym w stopniu  $d = 1$ ;  $Y_t \sim I(1)$ .

## 5. MODELE AR, MA, ARMA I ARIMA

Często w przypadku modeli szeregów czasowych zakłada się, że zjawiska ekonomiczne, będące procesami stochastycznymi, są generowane przez siebie same.

### Modele autoregresji

Do opisu procesów stacjonarnych przydatne okazują się modele autoregresji **AR**. W modelu autoregresyjnym bieżąca wartość zmiennej objaśnianej wyrażona jest przez skończoną kombinację jej wartości poprzednich.

Przyjmijmy, że w kolejnych chwilach czasu obserwuje się szereg czasowy postaci:  $y_t, y_{t-1}, y_{t-2}, \dots, y_T$ , będący realizacją **stacjonarnego procesu stochastycznego** o średniej  $\mu$ . Oznaczając przez  $\tilde{y}_t = y_t - \mu$  odchylenia od wartości średniej procesu, otrzymuje się szereg czasowy odchyleń czyli:  $\tilde{y}_t, \tilde{y}_{t-1}, \tilde{y}_{t-2}, \dots$ , do opisu którego zastosować można **model autoregresji rzędu  $p$**  oznaczony jako **AR( $p$ )**:

$$\tilde{y}_t = \varphi_1 \tilde{y}_{t-1} + \varphi_2 \tilde{y}_{t-2} + \dots + \varphi_p \tilde{y}_{t-p} + \varepsilon_t,$$

gdzie:

$\varepsilon_t$  - zmienna losowa o rozkładzie normalnym o średniej zero i stałej wariancji  $\sigma^2$ ,

$\varphi_j$  - parametry modelu.

Określmy operator przesunięcia wstecz  $B$  jako:  $B^s \tilde{y}_t = \tilde{y}_{t-s}$

Za jego pomocą definiujemy operator autoregresji rzędu  $p$ :  $\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p$

Model autoregresji można teraz zapisać:  $\varphi(B) \tilde{y}_t = (1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p) \tilde{y}_t = \varepsilon_t$

Proces autoregresji rzędu pierwszego AR(1):  $\tilde{y}_t = \varphi_1 \tilde{y}_{t-1} + \varepsilon_t$

Proces autoregresji rzędu drugiego AR(2):  $\tilde{y}_t = \varphi_1 \tilde{y}_{t-1} + \varphi_2 \tilde{y}_{t-2} + \varepsilon_t$

### Modele średniej ruchomej

W modelu średniej ruchomej **MA** (*moving average*) zakłada się, że  $\tilde{y}_t$  zależy liniowo od skończonej liczby  $q$  poprzednich wartości  $\varepsilon_t$ .

Przykładowo, dla procesu AR(1):

$$\begin{aligned} \tilde{y}_t &= \tilde{\varphi} \tilde{y}_{t-1} + \varepsilon_t = \varphi(\tilde{\varphi} \tilde{y}_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \\ &= \varepsilon_t + \varphi \varepsilon_{t-1} + \varphi^2 (\tilde{\varphi} \tilde{y}_{t-3} + \varepsilon_{t-2}) = \\ &= \varepsilon_t + \varphi \varepsilon_{t-1} + \varphi^2 \varepsilon_{t-2} + \dots + \varphi^q \varepsilon_{t-q} + \dots \end{aligned}$$

**Proces średniej ruchomej rzędu  $q$** , co oznacza się przez **MA( $q$ )**, jest postaci:

$$\tilde{y}_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

Definiując operator średniej ruchomej rzędu  $q$  jako:  $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$

model MA przekształca się do postaci:  $\tilde{y}_t = \theta(B)\varepsilon_t$ .

Proces średniej ruchomej rzędu pierwszego MA(1):  $\tilde{y}_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$

Proces średniej ruchomej rzędu drugiego MA(2):  $\tilde{y}_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$

### Modele ARMA

W celu osiągnięcia lepszego dopasowania modelu do rzeczywistego szeregu czasowego uzasadnione jest włączenie do modelu zarówno członów autoregresji AR( $p$ ), jak i średniej ruchomej MA( $q$ ).

Mieszany **model autoregresji i średniej ruchomej ARMA( $p, q$ )**:

$$\tilde{y}_t = \varphi_1 \tilde{y}_{t-1} + \varphi_2 \tilde{y}_{t-2} + \dots + \varphi_p \tilde{y}_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

lub inaczej:  $\varphi(B)\tilde{y}_t = \theta(B)\varepsilon_t$

Proces ARMA(1,1) ma postać:  $\tilde{y}_t - \varphi_1 \tilde{y}_{t-1} = \varepsilon_t - \theta_1 \varepsilon_{t-1}$

Proces budowy modeli ARMA jest procesem iteracyjnym podzielonym na trzy etapy: identyfikację, która ma na celu ustalenie rzędów autoregresji i średniej ruchomej, a w dalszym postępowaniu: estymację i weryfikację.

Celem określenia rzędów  $p$  i  $q$  dla procesów autoregresji i średniej ruchomej bada się postać estymowanych **funkcji autokorelacji i autokorelacji cząstkowej**.

Jeżeli proces jest **AR( $p$ )**, to funkcja autokorelacji maleje łagodnie, a funkcja autokorelacji cząstkowej urywa się nagle po opóźnieniu  $p$ . Jeżeli sytuacja jest odwrotna, to znaczy funkcja autokorelacji cząstkowej procesu maleje w sposób łagodny, natomiast funkcja autokorelacji urywa się po opóźnieniu  $q$ , to proces można określić jako proces **MA( $q$ )**. W przypadku procesu mieszanego nie stwierdza się, że funkcja autokorelacji czy autokorelacji cząstkowej urywa się nagle. Funkcja autokorelacji ma w tym przypadku tę właściwość, że po pierwszych ( $p - q$ ) opóźnieniach staje się zanikającą kombinacją funkcji wykładniczych i sinusoid tłumionych. Podobnie dzieje się z funkcją autokorelacji cząstkowej, z tym że funkcja zaczyna zauważalnie zanikać po pierwszych ( $p - q$ ) opóźnieniach. Na tej podstawie można określić rzędy  $p$  autoregresji i  $q$  średniej ruchomej procesu mieszanego **ARMA( $p, q$ )**.

### Modele procesów niestacjonarnych

Wiele zjawisk gospodarczych ma charakter niestacjonarny, co oznacza, że nie można wyróżnić w ich przebiegu wartości średniej, wokół której zjawisko oscyluje.

Niemniej jednak często własności szeregów niestacjonarnych są w pewien sposób jednorodne. Może się okazać, że mimo iż poziom, względem którego zachodzą wahania, zmienia się w czasie, to ogólne zachowanie się szeregu, po uwzględnieniu różnic w poziomie, będzie podobne.

Celem osiągnięcia stacjonarności badanego procesu, o którym zakładamy, że jest zintegrowany w stopniu  $d$ , obliczamy  $d$ -krotne różnice szeregu  $y_t$  ( $z_t = \Delta^d y_t$ ).

Do opisu szeregu  $z_t$  wykorzystuje się mieszany model autoregresji i średniej ruchomej. Model ten nazywany jest **scalkowanym procesem autoregresji i średniej ruchomej rzędu ( $p, d, q$ )**, w skrócie **ARIMA( $p, d, q$ )**.

Np. model ARIMA(2, 1, 2) jest postaci:  $\Delta y_t = \varphi_1 \Delta y_{t-1} + \varphi_2 \Delta y_{t-2} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$

Do uzyskania stacjonarnego szeregu  $z_t$  wystarcza z reguły wyliczenie pierwszych lub drugich różnic.



## 6. INNE MODELE ZE ZMIENNYMI OPÓŹNIONYMI

Zmienne opóźnione - zarówno objaśniające jak i objaśniane - występują bardzo często w modelach ekonometrycznych. Przyczynami ich występowania są:

- czynniki psychologiczne (przyzwyczajenia, bezwładność zachowań, koszt dostosowania),
- czynniki technologiczne (zmiany technologii wymagające czasu),
- czynniki instytucjonalno-prawne (zawarte umowy, zobowiązania, kontrakty, lokaty).

### Model z rozkładem opóźnień DL

(skończonym lub nieskończonym):

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \dots + \beta_k x_{t-k} + \varepsilon_t$$

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \varepsilon_t$$

W modelu tym wartości zmiennej objaśnianej są opisywane przez zmienną objaśniającą i jej wartości opóźnione.

Współczynniki  $\beta_0, \beta_1, \beta_2, \dots$  mierzą wpływ kolejnych opóźnień zmiennej objaśnianej na wielkość zmiennej objaśnianej.

Współczynnik  $\beta_0$  nazywamy **mnożnikiem krótkookresowym**. Określa on średnią zmianę zmiennej objaśnianej  $Y_t$ , wynikającą z jednostkowej zmiany zmiennej objaśnianej  $X_t$  dokonanej w tym samym czasie co obserwowana zmiana  $Y_t$ .

Sumę wszystkich współczynników  $\beta_i$  nazywamy **mnożnikiem długookresowym** i interpretujemy jako zmianę zmiennej objaśnianej  $Y_t$  pod wpływem długookresowej jednostkowej zmiany – jednostkowego przyrostu wszystkich wartości bieżących i opóźnionych – zmiennej objaśnianej.

### Model Koycka z nieskończonym rozkładem opóźnień

Przyjmijmy, że w modelu występuje nieskończony rozkład opóźnień:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \varepsilon_t$$

Dodatkowo zakładamy, że parametry  $\beta_k$  maleją geometrycznie, czyli, że zachodzi:

$$\beta_k = \beta_0 \lambda^k, \quad k = 0, 1, 2, \dots, \quad 0 < \lambda < 1.$$

Parametr  $\lambda$  nazywamy stopą zaniku rozkładu opóźnień.

Powyższy model możemy poddać tzw. transformacji Koycka. Zapiszmy model w następujących dwóch postaciach:

$$y_t = \alpha + \beta_0 x_t + \beta_0 \lambda^1 x_{t-1} + \beta_0 \lambda^2 x_{t-2} + \dots + \varepsilon_t$$

$$\lambda y_{t-1} = \lambda \alpha + \lambda \beta_0 x_{t-1} + \beta_0 \lambda^2 x_{t-2} + \beta_0 \lambda^3 x_{t-3} + \dots + \varepsilon_{t-1}$$

Odejmując równania stronami otrzymujemy:

$$y_t - \lambda y_{t-1} = \alpha(1 - \lambda) + \beta_0 x_t + \varepsilon_t - \lambda \varepsilon_{t-1}$$

Ostatecznie  $y_t = \sigma_0 + \lambda y_{t-1} + \beta_0 x_t + v_t$ , gdzie:  $\sigma_0 = \alpha(1 - \lambda)$  oraz  $v_t = \varepsilon_t - \lambda \varepsilon_{t-1}$ .

Transformacja Koycka sprowadziła model z nieskończonym rozkładem opóźnień do modelu z opóźnioną o jeden okres zmienną objaśnianą, czyli do modelu autoregresyjnego.

Mnożnik długookresowy w modelu Koycka wynosi:  $\beta_0 + \beta_0 \lambda + \beta_0 \lambda^2 + \beta_0 \lambda^3 + \dots = \frac{\beta_0}{1 - \lambda}$ .

### Model autoregresyjny z rozkładem opóźnień ADL

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_m y_{t-m} + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \beta_k x_{t-k} + \varepsilon_t$$

## 7. KOINTEGRACJA SZEREGÓW CZASOWYCH

Jeżeli w modelu  $y_t = \alpha + \beta x_t + \varepsilon_t$  zmienne  $X_t$  oraz  $Y_t$  są niestacjonarne, ale są zintegrowane w tym samym stopniu  $d$ , to może mieć miejsce **regresja pozorna** lub **kointegracja**.

**Regresja pozorna** miałaby miejsce, gdyby  $X_t$  i  $Y_t$  posiadały pierwiastki jednostkowe, a jednocześnie wektor składnika losowego nie był stacjonarny ( $\varepsilon_t \sim I(d)$ , gdzie  $d \geq 1$ ).

W wypadku regresji pozornej szacowanie modelu metodą KMNK może prowadzić do uzyskania wyników, które nie będą dawały dobrego ekonometrycznego modelu dla opisu jakiegoś zjawiska. Często może to być trudne do wcześniejszego wykrycia bez analizy pierwiastków jednostkowych, gdyż inne testy np. na istotność parametrów lub autokorelację składnika losowego mogą dawać dobre wyniki. Również współczynnik determinacji liniowej  $R^2$  może być wysoki.

Jedną z metod rozpoznawania, czy zachodzi regresja pozorna, jest test Engle-Grangera. Po przekształceniu równania  $y_t = \alpha + \beta x_t + \varepsilon_t$  do postaci  $\varepsilon_t = y_t - \beta x_t$  należy sprawdzić, czy szereg  $\varepsilon_t \sim I(0)$ . Jeżeli szereg  $\varepsilon_t \sim I(0)$ , wówczas nie mamy do czynienia z regresją pozorną.

**Kointegracja** między zmiennymi występuje właśnie, gdy składnik losowy modelu  $y_t = \alpha + \beta x_t + \varepsilon_t$  jest stacjonarny. Kointegrację rozumiemy jako występowanie dynamicznej zależności w czasie między zmiennymi  $X_t$  i  $Y_t$ .

W oparciu o zależność kointegracyjną dla zmiennych  $X_t$  i  $Y_t$  można wyznaczyć dla nich postać tzw. równowagi długookresowej. Różnica między równowagą długo- i krótkookresową może być traktowana jako miara bieżącego niedopasowania omawianego zjawiska. Można ją modelować i interpretować z wykorzystaniem modeli korekty błędem – ECM.

### Model korekty błędem – ECM

Koncepcja konstruowania modeli ECM związana jest z osobami Engle'a i Grangera, którzy poszukiwali zależności mogących wyznaczać i rozpoznawać różnice w zależnościach krótko- i długo-okresowych.

Model w równowadze długookresowej można zapisać jako:

$$y_t = \alpha + \beta x_t + \varepsilon_t, \quad \text{gdzie} \quad \varepsilon_t = ECM_t.$$

Wtedy model korekty błędem przyjmuje postać:

$$\Delta y_t = \alpha + \beta \Delta x_t + \gamma ECM_{t-1} + u_t.$$

Wartość parametru  $\gamma < 0$ , ponieważ przeciwnie mielibyśmy do czynienia z modelem "eksplodującym" w czasie.

Za pomocą modelu ECM możemy analizować zmiany w przyrostach  $\Delta y_t$  jako wynik zmian dla  $\Delta x_t$  oraz  $ECM_{t-1}$ ; czyli jak przyrost pewnego zjawiska zależy od przyrostu zjawiska, które go objaśnia oraz błędu w dopasowaniu modelu równowagi długookresowej do danych empirycznych.

Gdy  $\Delta x_t \rightarrow 0$  to  $ECM_{t-1}$  może mieć szczególnie duży wpływ na  $\Delta y_t$ . Ów błąd może być traktowany jako bieżąca miara niedoszacowania (gdy  $ECM_{t-1} < 0$ ) lub przeszacowania (dla  $ECM_{t-1} > 0$ ) dla oczekiwanej równowagi długookresowej względem danych empirycznych. Parametr  $\gamma$  pokazuje, jaka część błędu jest korygowana przez model w kolejnym okresie. Jeżeli jest ona zbyt mała, to można podać pod wątpliwość sens budowy modelu ECM.



## V. PROGNOZOWANIE EKONOMETRYCZNE

1. Prognoza
2. Błędy prognozy
3. Prognozowanie na podstawie modeli tendencji rozwojowej
4. Inne metody prognozowania

### 1. PROGNOZA

Finalnym etapem analizy ekonometrycznej jest wykorzystanie oszacowanego i zweryfikowanego modelu do prognozy zmiennej objaśnianej.

**Prognozowanie ekonometryczne** oznacza wnioskowanie o przyszłości za pomocą modelu, który został poddany wszechstronnej weryfikacji w poprzednich etapach. **Prognoza** to konkretny wynik tego wnioskowania.

Z modelu prognozuje się przyszłe wartości zmiennej objaśnianej (zmiennej prognozowanej) na podstawie danych przyszłych wartości zmiennych objaśniających.

Termin „prognozowanie” jest najczęściej używany z myślą o przewidywaniu przyszłości, jednak zasady predykcji ekonometrycznej mogą być stosowane i dla danych przekrojowych.

Przy prognozowaniu na podstawie modelu ekonometrycznego powinny być spełnione następujące **założenia**:

1. Ma miejsce stabilność relacji ujętych w modelu, tzn. postać analityczna, parametry i zbiór zmiennych objaśniających nie zmieniają się.
2. Ma miejsce stabilność rozkładu odchyłeń losowych modelu.
3. Znane są wartości zmiennych objaśniających, na podstawie których buduje się prognozę.
4. Dopuszczalna jest ekstrapolacja modelu poza próbę statystyczną.

Prognozę konkretnej zmiennej stanowi liczba (**prognoza punktowa**) lub przedział liczbowy, który z określonym prawdopodobieństwem zawiera wartość zmiennej prognozowanej (**prognoza przedziałowa**).

Niech oszacowany model ma postać  $\hat{y}_i = b_0 + b_1x_{1i} + \dots + b_kx_{ki}$ ,  $i=1, \dots, n$ .

Oznaczmy przez  $y_m^P$  wartość **prognozy punktowej** zmiennej  $Y$  na moment  $m$ . Przyjmując za wektor wartości zmiennych objaśniających  $\mathbf{x}_m^{*T} = [1 \ x_{1m}^* \ \dots \ x_{km}^*]^T$ , prognozę definiujemy następująco:

$$y_m^P = b_0 + b_1x_{1m}^* + \dots + b_kx_{km}^* = \mathbf{x}_m^{*T} \mathbf{b}.$$

Sposoby określenia przyszłych wartości zmiennych objaśniających mogą być następujące:

- analiza wartości przeszłych i przyjęcie dla okresu prognozowanego np. średniej wartości lub różnych wariantów wartości zmiennej (symulacja ekonometryczna),
- zbudowanie odrębnego modelu objaśniającego przebieg interesującej zmiennej,
- wykorzystanie metod ekstrapolacji szeregów czasowych.

#### Przykład 7.

Na podstawie 6 obserwacji oszacowano następujący model:

$$\hat{y}_i = 0,43302 + 0,16798 \cdot x_i,$$

gdzie  $y_i$  – dostawy serów w tys. ton,  $x_i$  – dostawy masła w tys. ton.

Jeśli chcemy prognozować, jakie byłyby dostawy serów w roku 2007 w przypadku dostaw masła w tym roku na poziomie 70 tys. ton, wyznaczmy prognozę punktową:

$$y_{2007}^P = 0,43302 + 0,16798 \cdot x_{2007}^* = 0,43302 + 0,16798 \cdot 70 = 12,19162$$

## 2. BŁĘDY PROGNOZY

Przy ocenie jakości wyznaczonej prognozy punktowej wykorzystuje się dwa typy błędów:

- oczekiwany błąd prognozy (**błąd ex ante**), nazywany błędem predykcji lub predyktora,
- zrealizowany błąd prognozy (**błąd ex post**), nazywany błędem prognozy.

Niech  $y_m^*$  oznacza rzeczywistą wartość zmiennej prognozowanej, natomiast  $y_m^P$  prognozę punktową dla tej zmiennej. Wtedy błąd prognozy jest równy  $y_m^* - y_m^P$ .

Błąd ten jest zmienną losową o wartości oczekiwanej równej zero i odchyleniu standardowym

$$D(y_m^* - y_m^P) = \sigma \sqrt{\mathbf{x}_m^{*T} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{x}_m^* + 1}.$$

**Średnim błędem predykcji ex ante** jest estymator odchylenia standardowego błędu prognozy wyrażający się wzorem:

$$v_m = S_e \sqrt{\mathbf{x}_m^{*T} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{x}_m^* + 1},$$

w którym

$S_e$  - ocena nieznanego odchylenia standardowego składnika losowego  $\sigma$  modelu,

$\mathbf{x}_m^*$  - wektor przyszłych wartości zmiennych objaśniających dla momentu  $m$ ,

$\mathbf{X}$  - macierz obserwacji na zmiennych objaśniających, na podstawie których oszacowany został model.

Błąd ex ante wyznaczany jest w momencie budowy prognozy i informuje, o ile oszacowana wartość zmiennej prognozowanej  $y_m^P$  odchyła się średnio od rzeczywistej wartości zmiennej prognozowanej  $y_m^*$ .

Do porównywania dokładności prognoz wyznaczonych dla różnych okresów czasu lub dla różnych zmiennych przydatny jest błąd względny predykcji postaci:  $wv_m = \frac{v_m}{y_m^P}$ .

Średni błąd predykcji ex ante wykorzystywany jest do wyznaczenia **prognozy przedziałowej** dla pojedynczej wartości  $y_m^*$ .

$(1-\alpha) \cdot 100$ -procentowy **przedział ufności dla prognozy** (inaczej **przedział predykcji**), jest postaci:

$$(y_m^P - t^* \cdot v_m; y_m^P + t^* \cdot v_m),$$

gdzie  $y_m^P$  - prognoza punktowa,

$t^*$  - wartość krytyczna z tablic dla rozkładu t-Studenta dla  $n-k-1$  stopni swobody i poziomu istotności  $\alpha$ ,

$v_m$  - średni błąd predykcji ex ante.

Z prawdopodobieństwem  $1-\alpha$  można twierdzić, że przedział powyższy pokryje rzeczywistą wartość zmiennej prognozowanej  $y_m^*$ .

**Błąd prognozy ex post** wyznacza się dopiero po dokonaniu pomiaru rzeczywistej wartości zmiennej prognozowanej. Jest on różnicą między zaobserwowaną wartością zmiennej prognozowanej  $y_m^*$  i wartością prognozy  $y_m^P$ :

$$e_m^* = y_m^* - y_m^P$$

Błąd względny prognozy ex post określony jest jako:  $we_m^* = \frac{y_m^* - y_m^P}{y_m^*}$

Do najczęściej stosowanych w praktyce innych miar dokładności prognoz ex post należą:

- średni błąd (*mean error*):



$$ME = \frac{1}{T} \sum_{m=1}^T (y_m^* - y_m^P)$$

- średni błąd bezwzględny (*mean absolute error*):

$$MAE = \frac{1}{T} \sum_{m=1}^T |y_m^* - y_m^P|$$

- średni absolutny błąd procentowy (*mean absolute percentage error*):

$$MAPE = \frac{1}{T} \sum_{m=1}^T \left| \frac{y_m^* - y_m^P}{y_m^*} \right| \cdot 100$$

- błąd średniokwadratowy (*mean squared error*):

$$MSE = \frac{1}{T} \sum_{m=1}^T (y_m^* - y_m^P)^2$$

- pierwiastek błędu średniokwadratowego (*root mean squared error*):

$$RMSE = \sqrt{\frac{1}{T} \sum_{m=1}^T (y_m^* - y_m^P)^2}$$

gdzie:

$y_m^*$  - wartość rzeczywista zmiennej prognozowanej,

$y_m^P$  - wyznaczona prognoza,

$T$  - horyzont prognozy.

### 3. PROGNOZOWANIE NA PODSTAWIE MODELI TENDENCJI ROZWOJOWEJ

Prognozowanie na podstawie modelu tendencji rozwojowej  $\hat{y}_t = f(t) + g_t(t)$  polega na prostej **ekstrapolacji funkcji trendu** (i ewentualnie funkcji  $g_t(t)$ ) przez podstawienie do modelu w miejsce zmiennej czasowej numeru momentu  $m$ , na który wyznacza się prognozę:

$$y_m^P = f(m) + g_t(m) \quad \text{dla} \quad m > n.$$

Przykładowo założmy, że mamy oszacowany model trendu liniowego postaci:  $\hat{y}_t = b_0 + b_1 t$ .

Wówczas prognozę punktową na moment  $m$ -ty wyznacza się jako:  $y_m^P = b_0 + b_1 m$ .

#### Przykład 8.

Zmieniającą się w ciągu 12 miesięcy cenę pewnego dobra opisuje model trendu:

$$\hat{y}_t = 30,08 + 0,55t.$$

Podstawiając  $t = 13$  wyznaczmy prognozę ceny na kolejny miesiąc:

$$y_{13}^P = 30,08 + 0,55 \cdot 13 = 37,23.$$

**Wyodrębnianie wahań sezonowych** przeprowadza się za pomocą tzw. **wskaźnika sezonowości**, który można wyznaczyć

$$\text{analitycznie:} \quad S_l = \frac{\sum_{i=0}^{T-1} y_{l+iL}}{\sum_{i=0}^{T-1} \hat{y}_{l+iL}} \quad \text{lub} \quad \text{mechanicznie:} \quad S_l = \frac{\sum_{i=0}^{T-1} y_{l+iL}}{\sum_{i=0}^{T-1} \bar{y}_{l+iL}}$$

gdzie:

$\hat{y}_t$  - wartości teoretyczne wyznaczone na podstawie oszacowanej funkcji trendu,

$\bar{y}_t$  - wartości średnich ruchomych obliczone dla kolejnych podokresów,

$S_l$  - surowe wskaźniki sezonowości,

$t = 1, 2, \dots, T, l = 1, 2, \dots, L$  (dla danych kwartalnych  $L=4$ , miesięcznych  $L=12$ ).

Suma surowych wskaźników sezonowości powinna być równa  $L$ , czyli:  $\sum_{l=1}^L S_l = L$ .

Jeśli ta relacja nie zachodzi, to należy wyznaczyć **współczynniki skorygowane** postaci:

$$S_l^v = v \cdot S_l$$

gdzie:  $v$  - współczynnik korygujący, który wyznacza się jako:  $v = \frac{L}{\sum_{l=1}^L S_l}$ .

**Absolutną wielkość odchyień sezonowych** wyznacza się z relacji:

$$g_l(t) = S_l^v \cdot \bar{y} - \bar{y},$$

gdzie:  $\bar{y}$  - średni poziom zjawiska  $y_t$ , wyznaczony dla wszystkich obserwacji  $t = 1, 2, \dots, T$ .

#### 4. INNE METODY PROGNOZOWANIA

Gdy w szeregu czasowym występuje składowa systematyczna w postaci stałego (przeciętnego) poziomu i wahań przypadkowe, do prognozowania używa się zwykle następujących metod:

- metodę naiwną,
- metody średniej ruchomej,
- model wygładzania wykładniczego.

Metody te umożliwiają konstrukcję prognoz krótkookresowych - na jeden kolejny okres.

**Metodę naiwną** można stosować jedynie przy niewielkich wahań przypadkowych. W metodzie tej konstruuje się prognozę zmiennej na moment  $t$  na poziomie zaobserwowanej wartości tej zmiennej w momencie  $t-1$ , czyli:

$$y_t^p = y_{t-1},$$

gdzie:  $y_t^p$  - prognoza zmiennej  $Y$  wyznaczona na chwilę  $t$ ,

$y_{t-1}$  - wartość zmiennej  $Y$  w chwili  $t-1$ .

**Metody średniej ruchomej** stosuje się do prognozowania, gdy poziom zmiennej prognozowanej jest prawie stały, z niewielkimi odchyleniami losowymi, a w szeregu czasowym nie występują inne składowe.

W metodach tych przyjmuje się, że wartość zmiennej prognozowanej w następnym okresie jest równa średniej arytmetycznej z  $k$  ostatnich wartości tej zmiennej.

Obliczanie prognozy na podstawie modelu **średniej ruchomej prostej** jest następujące:

$$y_t^p = \frac{1}{k} \sum_{i=t-k}^{t-1} y_i,$$

gdzie:  $y_t^p$  - prognoza zmiennej  $Y$  wyznaczona na chwilę  $t$ ,

$y_i$  - wartość zmiennej  $Y$  w chwili  $i$ .

Liczba wyrazów średniej ruchomej  $k$  (stała wygładzania) jest określana przez prognostę. Jej wybór może być oparty na minimalizacji wartości błędu prognozy ex post  $MAE$ .

**Wygładzanie wykładnicze** polega na tym, że szereg czasowy zmiennej prognozowanej wygładza się za pomocą ważonej średniej ruchomej, przy czym wagi są określane wykładniczo. Nowsze obserwacje, które zawierają bardziej aktualne informacje o prognozowanym zjawisku, dostają relatywnie większe wagi niż obserwacje starsze



**(postulat postarzania informacji).** Wygładzanie wykładnicze może być oparte na różnych modelach, w zależności od rodzaju składowych szeregu czasowego.

Prosty model wygładzania wykładniczego może być stosowany do prognozowania w przypadku, gdy w szeregu czasowym występuje składowa systematyczna w postaci stałego (przeciętnego) poziomu i wahań przypadkowe. Prognozę oblicza się wtedy następująco:

$$y_t^P = \alpha y_{t-1} + (1 - \alpha) y_{t-1}^P, \quad \alpha \in (0, 1),$$

gdzie:

$y_t^P, y_{t-1}^P$  - prognozy zmiennej  $Y$  wyznaczone na chwile  $t, t-1$ ,

$y_{t-1}$  - wartość zmiennej  $Y$  w chwili  $t-1$ ,

$\alpha$  - parametr wygładzania.

Inne popularne modele wyrównania wykładniczego to model Browna, model Holta oraz model z dwoma parametrami wyrównywania.

## VI. MODELE ZE ZMIENNYMI JAKOŚCIOWYMI

1. Zmienne jakościowe w charakterze zmiennych objaśniających
2. Wykorzystanie zmiennych jakościowych do analizy danych sezonowych
3. Zmienne jakościowe w charakterze zmiennych objaśnianych

### 1. ZMIENNE JAKOŚCIOWE W CHARAKTERZE ZMIENNYCH OBJAŚNIAJĄCYCH

Aby takie zmienne jakościowe, jak płeć, wykształcenie, stan cywilny, mogły być uwzględnione w modelu ekonometrycznym, należy dokonać ich kwantyfikacji za pomocą zmiennych zero-jedynkowych.

Jeśli zmienna jakościowa odnosi się do 2 możliwych wariantów  $A$  i  $B$ , to można zdefiniować odpowiadającą jej zmienną zero-jedynkową jako:

$$z_{1i} = \begin{cases} 1, & \text{gdy obserwacja reprezentuje wariant } A, \\ 0, & \text{gdy obserwacja reprezentuje wariant } B. \end{cases}$$

Wówczas w modelu

$$\hat{y}_i = c_0 + a_1 x_{1i} + c_1 z_{1i}$$

$c_1$  mierzy średni wpływ na zmienną objaśnianą wariantu  $A$  w odniesieniu do wariantu  $B$  (przy takim samym poziomie zmiennej ilościowej  $X_1$ ).

#### Przykład 9.

Oszacowano model:  $\hat{y}_i = 1,2 + 0,21x_{1i} - 0,5z_{1i}$ , w którym  $Y$  - miesięczne spożycie słodczy w kg,  $X_1$  - miesięczne dochody w tys. zł,  $Z_1$  - zmienna zero-jedynkowa, przyjmującą wartość 1, jeśli badana osoba jest mężczyzną oraz wartość 0, gdy badaną osobą jest kobieta.

Po wyeliminowaniu wpływu płci wzrost miesięcznych dochodów o 1 tysiąc złotych powoduje wzrost miesięcznego spożycia słodczy średnio o 0,21 kg. Z kolei przy takich samych dochodach mężczyźni spożywają miesięcznie średnio o 0,5 kg słodczy mniej niż kobiety.

W przypadku, gdy zmienna jakościowa odnosi się do 3 możliwych wariantów  $A$ ,  $B$  i  $C$ , to zmienne zero-jedynkowe określa się jako:

$$z_{1i} = \begin{cases} 1, & \text{gdy } i\text{-ta obserwacja reprezentuje wariant } A, \\ 0, & \text{w przeciwnych przypadkach,} \end{cases}$$

oraz

$$z_{2i} = \begin{cases} 1, & \text{gdy } i\text{-ta obserwacja reprezentuje wariant } B, \\ 0, & \text{w przeciwnych przypadkach.} \end{cases}$$

W modelu ze stałą nie należy uwzględniać zmiennej odnoszącej się do trzeciego wariantu. Gdyby bowiem dołączona została zmienna  $Z_3$ :

$$z_{3i} = \begin{cases} 1, & \text{gdy } i\text{-ta obserwacja reprezentuje wariant } C, \\ 0, & \text{w przeciwnych przypadkach,} \end{cases}$$

wówczas między stałą i zmiennymi  $Z_1$ ,  $Z_2$ ,  $Z_3$ , zachodziłaby zależność liniowa, zatem nie byłoby możliwe jednoznaczne oszacowanie parametrów modelu:

$$y_i = \beta_0 + \alpha_1 x_{1i} + \beta_1 z_{1i} + \beta_2 z_{2i} + \beta_3 z_{3i} + \varepsilon_i.$$

W modelu ze stałą, liczba zmiennych reprezentujących zmienną jakościową musi być o jeden mniejsza od liczby wariantów.



Jeśli w modelu  $y_i = \beta_0 + \alpha_1 x_{1i} + \beta_1 z_{1i} + \beta_2 z_{2i} + \beta_3 z_{3i} + \varepsilon_i$  usuniemy zmienną  $Z_3$  odpowiadającą wariantowi C, wówczas uzyskamy model:

$$y_i = \beta_0 + \alpha_1 x_{1i} + \beta_1 z_{1i} + \beta_2 z_{2i} + \beta_3 (1 - z_{1i} - z_{2i}) + \varepsilon_i$$

$$y_i = (\beta_0 + \beta_3) + \alpha_1 x_{1i} + (\beta_1 - \beta_3) z_{1i} + (\beta_2 - \beta_3) z_{2i} + \varepsilon_i.$$

Ocena parametru występująca w modelu przy zmiennej reprezentującej dany wariant mierzy średni wpływ na zmienną objaśnianą tego wariantu odniesiony do wpływu wariantu pominiętego.

## 2. WYKORZYSTANIE ZMIENNYCH JAKOŚCIOWYCH DO ANALIZY DANYCH SEZONOWYCH

Zmienne jakościowe mogą być także wykorzystane do analizy danych sezonowych. Jeżeli w badanym zjawisku występują wahania sezonowe, które są liniowe i stałe dla poszczególnych kwartałów, to można opisać je w następujący sposób:

$$y_i = \beta_0 + \sum_{l=1}^4 \beta_l z_{li} + \sum_{j=1}^k \alpha_j x_{ji} + \varepsilon_i,$$

gdzie:

$Z_l$  – zmienne zero-jedynkowe przyjmujące wartość jeden w  $l$ -tym kwartale każdego roku i wartość zero w pozostałych kwartałach,  $l=1,2,3,4$ ,

$X_j$  – pozostałe zmienne objaśniające,  $j=1,2, \dots, k$ .

Ponieważ liczba zmiennych reprezentujących zmienną jakościową musi być o jeden mniejsza od liczby wariantów, podstawiamy  $z_{4i} = 1 - z_{1i} - z_{2i} - z_{3i}$  i otrzymujemy

$$y_i = \beta_0 + \beta_1 z_{1i} + \beta_2 z_{2i} + \beta_3 z_{3i} + \beta_4 (1 - z_{1i} - z_{2i} - z_{3i}) + \sum_{j=1}^k \alpha_j x_{ji} + \varepsilon_i.$$

Ostatecznie:

$$y_i = \gamma_0 + \sum_{l=1}^3 \gamma_l z_{li} + \sum_{j=1}^k \alpha_j x_{ji} + \varepsilon_i, \quad i=1,2,\dots,n,$$

gdzie  $\gamma_0 = \beta_0 + \beta_4$ ,  $\gamma_1 = \beta_1 - \beta_4$ ,  $\gamma_2 = \beta_2 - \beta_4$ ,  $\gamma_3 = \beta_3 - \beta_4$ .

Oceny  $c_l$  parametrów  $\gamma_l$  informują o średniej wielkości efektów sezonowych występujących w kwartale o numerze  $l$ ,  $l=1,2,3$ , w stosunku do wielkości zmiennej objaśnianej w IV kwartale. Jeżeli chcemy uzyskać informacje o odchyleniach od średniego kwartalnego poziomu zmiennej objaśnianej, to należy wyznaczyć oceny parametrów  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$  i  $\beta_4$ .

Ponieważ  $\gamma_0 = \beta_0 + \beta_4$ ,  $\gamma_l = \beta_l - \beta_4$ ,  $l=1,2,3$ , oraz  $\beta_4 = -\beta_1 - \beta_2 - \beta_3$ , to rozwiązania takiego układu 5 równań z 5 niewiadomymi  $\beta_l$ ,  $l=0,1,2,3,4$  są następujące:

$$\begin{aligned} \beta_0 &= \gamma_0 + \frac{1}{4}(\gamma_1 + \gamma_2 + \gamma_3) & \beta_1 &= \gamma_1 - \frac{1}{4}(\gamma_1 + \gamma_2 + \gamma_3) \\ \beta_2 &= \gamma_2 - \frac{1}{4}(\gamma_1 + \gamma_2 + \gamma_3) & \beta_3 &= \gamma_3 - \frac{1}{4}(\gamma_1 + \gamma_2 + \gamma_3) \\ \beta_4 &= -\frac{1}{4}(\gamma_1 + \gamma_2 + \gamma_3). \end{aligned}$$

### 3. ZMIENNE JAKOŚCIOWE W CHARAKTERZE ZMIENNYCH OBJAŚNIANYCH

Za pomocą modelu ekonometrycznego możemy opisywać zachowanie jednostek, wykorzystując przy tym zmienne jakościowe, najczęściej o charakterze zmiennych zero-jedynkowych. W wypadku, gdy zachodzi konieczność uwzględnienia w modelu ekonometrycznym zmiennej objaśnianej, przyjmującej wartość zero lub jeden (np. decyzja o zakupie samochodu), do opisu zjawiska proponuje się wykorzystanie liniowego modelu prawdopodobieństwa, modelu logitowego lub modelu probitowego.

**Liniowy model prawdopodobieństwa** ma postać:

$$y_i^* = \beta_0 + \beta_1 x_{1i} + \dots + \beta_{ki} + \varepsilon_i, \quad i = 1, 2, \dots, n$$

gdzie  $y_i^*$  to zmienna ukryta oraz

$$y_i = \begin{cases} 1, & \text{dla } y_i^* > 0 \\ 0, & \text{dla } y_i^* \leq 0 \end{cases}$$

**Model logitowy** to model następujący:

$$y_i^* = \ln\left(\frac{P_i}{1-P_i}\right) = \beta_0 + \beta_1 x_{1i} + \dots + \beta_{ki} + \varepsilon_i, \quad i = 1, 2, \dots, n,$$

gdzie  $y_i^*$  nazywany jest **logitem**, a  $P_i$  oznacza prawdopodobieństwo zmiennej zależnej  $y_i$  wyznaczone na podstawie rozkładu logistycznego i przyjmujące wartość z przedziału (0,1).

Wartość  $\hat{P}_i$  wyznaczamy ze wzoru:

$$\hat{P}_i = \frac{1}{1 + e^{-\hat{y}_i}} = \frac{1}{1 + e^{-(b_0 + b_1 x_{1i} + \dots + b_k x_{ki})}}.$$

Gdy  $y_i^* \rightarrow \infty$ , to  $P_i \rightarrow 1$ ; gdy  $y_i^* \rightarrow -\infty$ , to  $P_i \rightarrow 0$ ; a gdy  $y_i^* = 0$ , to  $P_i = 0,5$ .

W **modelu probitowym** wartość  $P_i$  wyznaczana jest z dystrybuanty rozkładu normalnego i wynosi:

$$P_i = F(b_0 + b_1 x_{1i} + \dots + b_k x_{ki}) = \int_{-\infty}^{b_0 + b_1 x_{1i} + \dots + b_k x_{ki}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$$

Między parametrami w modelach logitowym i probitowym zachodzi:  $\beta_{\text{logit}} = 1,6 \cdot \beta_{\text{probit}}$ .

Powyżej opisane modele szacuje się na podstawie danych indywidualnych przekrojowych. Zmienną objaśnianą jest w nich zmienna zero-jedynkowa. Przedstawia ona najczęściej wynik decyzji ekonomicznej, np. zakup samochodu, udzielenie kredytu, upadłość firmy.

**Prognozowanie zmiennej jakościowej** zero-jedynkowej polega na wyznaczeniu stanu jeden lub zero. W wypadku modelu logitowego dla odpowiednich wartości teoretycznych  $\hat{P}_i$  prognozujemy:

$$\hat{P}_i > 0,5 \Rightarrow \hat{y}_i = 1 \quad \text{lub} \quad \hat{P}_i \leq 0,5 \Rightarrow \hat{y}_i = 0.$$



Sposobem **oceny modelu logitowego** jest przedstawienie wyników prognozy dla jednostek z próby w formie **tablicy trafień**:

Faktyczne	Przewidywane		Razem
	Y=1	Y=0	
Y=1	$n_{11}$	$n_{10}$	$n_{1.}$
Y=0	$n_{01}$	$n_{00}$	$n_{0.}$
Razem	$n_{.1}$	$n_{.0}$	$n$

Procentowa trafność prognozowania jest następująca:

$$\text{łącznie: } \frac{n_{11} + n_{00}}{n}, \quad \text{dla } Y_i = 1 : \frac{n_{11}}{n_{1.}}, \quad \text{dla } Y_i = 0 : \frac{n_{00}}{n_{0.}}.$$

Trafność prognoz oddaje również iloraz szans postaci:  $IS = \frac{n_{11} \cdot n_{00}}{n_{01} \cdot n_{10}}.$

## VII. MODELE NIELINIOWE

1. Klasyfikacja modeli nieliniowych
2. Typy modeli nieliniowych stosowane w badaniach ekonomicznych
3. Metody estymacji modeli nieliniowych

### 1. KLASYFIKACJA MODELI NIELINIOWYCH

W ekonometrii często wykorzystujemy modele liniowe, gdyż mają prostą interpretację ekonomiczną, są łatwe w estymacji, łatwa jest ich statystyczna weryfikacja, często są wystarczająco dobrym przybliżeniem rzeczywistości ekonomicznej.

W jednorównaniowym **modelu liniowym**:

$$y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki} + \varepsilon_i, \quad i=1,2,\dots,n,$$

zmiennie objaśniające  $X_j$ , jak i parametry strukturalne  $\beta_j$ , pojawiają się w sposób liniowy.

Jednak w badaniach ekonomicznych pomiędzy zmiennymi często występują zależności nieliniowe, np. zależność między nakładami materiałowymi a wielkością produkcji końcowej.

W takich sytuacjach do opisu należy stosować odpowiedni model nieliniowy.

Ogólna postać **modelu nieliniowego** dla wielu zmiennych z addytywnym składnikiem losowym:

$$y_i = f(x_{1i}, \dots, x_{ki}, \theta) + \varepsilon_i,$$

gdzie  $f$  – dowolna funkcja nieliniowa,  $\theta$  – wektor parametrów strukturalnych.

Generalnie modele ekonometryczne można **podzielić na**:

a. liniowe,

b. nieliniowe:

- ze względu na zmiennie objaśniające, np.  $y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i$ ,
- ze względu na parametry, np.  $y_i = \beta_0 + \beta_1 x_{1i} + \beta_1^2 x_{2i} + \varepsilon_i$ ,
- ze względu na zmiennie objaśniające i parametry, ale sprowadzalne do modeli liniowych względem parametrów, np.  $y_i = \beta_0 \cdot x_{1i}^{\beta_1} \cdot x_{2i}^{\beta_2} \cdot \varepsilon_i$ ,
- ściśle nieliniowe, niesprowadzalne do postaci liniowej, np.  $y_i = \beta_1 + \beta_2 e^{\beta_3 \cdot x_{1i}} + \varepsilon_i$ .

### 2. TYPY MODELI NIELINIOWYCH STOSOWANE W BADANIACH EKONOMICZNYCH

**Modele liniowe ze względu na parametry**

- **Model wielomianowy**  $l$  – tego stopnia:  $\hat{y}_i = b_0 + b_1 x_{1i} + b_2 x_{1i}^2 + \dots + b_k x_{ki}^l$

Model ten, a szczególnie wielomian drugiego stopnia, bardzo często jest stosowany w badaniach ekonomicznych do wyjaśniania przyrostów wielkości produkcji w zależności od poziomu nakładów materiałowych. Może on również reprezentować funkcję przeciętnego kosztu w zależności od wielkości produkcji.

- **Model logarytmiczny**:  $\hat{y}_i = b_0 + b_1 \ln x_{1i} + b_2 \ln x_{2i} + \dots + b_k \ln x_{ki}$

W modelu tym wartości zmiennych objaśniających muszą być dodatnie. Gdy  $b_j > 0$ , to przyrostowi zmiennej objaśniającej  $X_j$  towarzyszą coraz mniejsze przyrosty zmiennej objaśnianej  $Y$  (np. przyrost plonu zbóż na jednostkę przyrostu nawożenia mineralnego),

a krańcowy przyrost  $Y$  na jednostkę przyrostu  $X_j$ , czyli  $\frac{\partial \hat{y}}{\partial x_j} = \frac{b_j}{x_j}$ , jest malejącą funkcją  $X_j$ .

- **Model hiperboliczny:**  $\hat{y}_i = b_0 + \frac{b_1}{x_{1i}} + \frac{b_2}{x_{2i}} + \dots + \frac{b_k}{x_{ki}}$

W przypadku modelu z jedną zmienną objaśniającą hiperbola ma asymptotę poziomą  $y = b_0$ , a jej przebieg zależy od znaku współczynnika  $b_1$ . Zależność jest malejąca, gdy współczynnik ten jest dodatni, w przeciwnym przypadku rosnąca.

- **Model z interakcjami:**  $\hat{y}_i = b_0 + b_1x_{1i} + b_2x_{2i} + b_3x_{1i}x_{2i}$

Modele z interakcjami uwzględniają to, że wpływ zmiany jednej zmiennej objaśniającej zależy od wartości przyjmowanej przez inne zmienne objaśniające. Pochodna cząstkowa  $Y$  po  $X_1$ ,  $\frac{\partial \hat{y}}{\partial x_1} = b_1 + b_3x_2$ , jest liniową funkcją  $X_2$ . Jeśli  $X_2$  będzie interpretowane jako czas, to krańcowy przyrost będzie liniową funkcją czasu.

### Modele linearyzowalne ze względu na parametry

- **Model potęgowy:**  $\hat{y}_i = b_0x_{1i}^{b_1}x_{2i}^{b_2}\dots x_{ki}^{b_k}$

Model ten bardzo często jest stosowany jako ekonometryczna funkcja produkcji Cobba–Douglasa lub jako ekonometryczna funkcja popytu. Jego podstawową własnością jest stała elastyczność zmiennej  $Y$  względem zmiennych  $X_j$ . Np. w modelu dla produkcji  $\hat{y}_i = 5,1 \cdot x_i^{0,8}$ , produkcja wzrasta średnio o 0,8% na 1% przyrostu nakładów.

- **Model wykładniczy:**  $\hat{y}_i = b_0 \cdot p^{b_1 \cdot x_{1i} + b_2 \cdot x_{2i} + \dots + b_k \cdot x_{ki}}$

Najczęściej używa się podstawy  $e$  lub  $10$ . Model wykładniczy wykorzystywany jest do określania średniego przyrostu względnego zmiennej objaśnianej na jednostkę przyrostu bezwzględnego zmiennej objaśniającej, przy średnim poziomie pozostałych zmiennych uwzględnionych w modelu. Np. w modelu dla produkcji  $\hat{y}_i = 26,2 \cdot e^{0,01 \cdot x_i}$ , produkcja wzrasta średnio o 1% na jednostkę bezwzględnego przyrostu nakładów.

- **Funkcja Törnquista I:**  $\hat{y}_i = \frac{ax_i}{b + x_i}$

Funkcja ta stosowana jest do badania zależności pomiędzy dochodem a wydatkami na dobra pierwszej potrzeby.

- **Funkcja Törnquista II:**  $\hat{y}_i = \frac{a(x_i - c)}{x_i + b}$

Funkcja ta stosowana jest do opisu zależności pomiędzy dochodem a wydatkami na dobra wyższego rzędu.

- **Funkcja Törnquista III:**  $\hat{y}_i = \frac{ax_i(x_i - c)}{x_i + b}$

Funkcja ta wykorzystywana jest do opisu zależności pomiędzy dochodem a wydatkami na dobra luksusowe.

- **Funkcja logistyczna:**  $\hat{y}_i = \frac{a}{1 + be^{-cx_i}}$

Funkcja ta należy do grupy funkcji s–kształtnych. Wartość parametru  $a$  określa poziom nasycenia, czyli maksymalny poziom wartości zmiennej  $Y$ . Dla  $x_i = \frac{\ln b}{c}$  punkt przegięcia.

- **Model Workinga:**  $\hat{y}_i = \exp(b_0 + \frac{b_1}{x_i})$

### Modele segmentowe

Jeżeli w analizie szeregów czasowych możemy wyodrębnić podokresy, w których zbiory obserwacji łatwo opisują się różnymi modelami ekonometrycznymi, to tworzymy model segmentowy. Element szeregu czasowego kończący jeden segment i jednocześnie zaczynający następny nazywamy punktem zwrotnym. Najprostszym przypadkiem jest model kawałkami liniowy; tu każdy segment to model trendu liniowego.

Brak jest jednolitych reguł wyboru postaci analitycznej modelu. Praktykowany jest wybór:

- na podstawie apriorycznej wiedzy o badanym zjawisku:
  - jeżeli przyrosty  $Y$  na jednostkę przyrostu  $X$  są stałe, to stosujemy model liniowy,
  - jeżeli elastyczność  $Y$  względem  $X$  jest stała, to stosujemy model potęgowy,
  - jeżeli  $Y$  charakteryzuje się malejącymi przyrostami na jednostkę przyrostu  $X$ , to stosujemy model logarytmiczny,
  - jeżeli mamy do czynienia ze stałymi względnymi przyrostami  $Y$  na jednostkę przyrostu bezwzględnego  $X$ , to stosujemy model wykładniczy;
- metodą prób i błędów - decyzję o wyborze jednej funkcji spośród wielu innych podejmujemy na podstawie dopasowania modelu do danych empirycznych, informacji o precyzji oszacowań parametrów oraz o rozkładzie reszt;
- na podstawie wykresu korelacyjnego dwóch zmiennych.

## 3. METODY ESTYMACJI MODELI NIELINIOWYCH

Nieliniowe modele mogą być szacowane na drodze:

- **linearyzacji**, a następnie wyznaczenia ocen parametrów transformowanego modelu klasyczną metodą najmniejszych kwadratów,
- **metod numerycznych estymacji nieliniowej** w wypadku modeli ściśle nieliniowych,
- **segmentowej estymacji** parametrów modeli nieliniowych o segmentach liniowych.

### Estymacja parametrów modeli nieliniowych poprzez linearyzację

Ogólna postać modelu liniowego ze względu na parametry:

$$g(y_i) = \beta_0 + \beta_1 f_1(x_{1i}, \dots, x_{ki}) + \dots + \beta_l f_l(x_{1i}, \dots, x_{ki}) + \varepsilon_i$$

Jeśli  $g(y_i) = y_i$ , to mówimy, że mamy do czynienia z modelem bezpośrednio liniowym. Szacowanie parametrów takiego modelu dokonywane jest KMNK przy odpowiednim przedefiniowaniu zmiennych.

Przykłady:

Funkcja	Postać strukturalna	Postać liniowa	Związek między zmiennymi
logarytmiczna	$\hat{y}_i = b_0 + b_1 \ln x_i$	$\hat{y}_i = b_0 + b_1 z_i$	$z_i = \ln x_i$
wielomianowa	$\hat{y}_i = b_0 + b_1 x_i + b_2 x_i^2$	$\hat{y}_i = b_0 + b_1 z_{1i} + b_2 z_{2i}$	$z_{1i} = x_i, z_{2i} = x_i^2$

Jeśli  $g(y_i)$  jest odwracalna, to model  $y_i = g^{-1}(\beta_0 + \beta_1 f_1(x_{1i}, \dots, x_{ki}) + \dots + \beta_l f_l(x_{1i}, \dots, x_{ki}) + \varepsilon_i)$  nazywamy modelem linearyzowanym. Model jest linearyzowalny, jeśli istnieje jednoznaczne jego przekształcenie, w wyniku którego otrzymujemy model liniowy. W tak przekształconym modelu oceny parametrów szacowane są KMNK.

Przykłady funkcji nieliniowych sprowadzalnych do postaci liniowej wzgl. parametrów:

Funkcja	Postać strukturalna	Postać liniowa	Związek między parametrami
potęgowa	$\hat{y}_i = a_0 x_i^{a_1}$	$\ln \hat{y}_i = b_0 + b_1 \ln x_i$	$a_0 = e^{b_0}, a_1 = b_1$
wykładnicza	$\hat{y}_i = a_0 \cdot e^{a_1 \cdot x_i}$	$\ln \hat{y}_i = b_0 + b_1 \cdot x_i$	$a_0 = e^{b_0}, a_1 = b_1$
wykładnicza z odwrotnością	$\hat{y}_i = a_0 e^{(a_1 \frac{1}{x_i})}$	$\ln \hat{y}_i = b_0 + b_1 \frac{1}{x_i}$	$a_0 = e^{b_0}, a_1 = b_1$
potęgowo-wykładnicza	$\hat{y}_i = a_0 x_i^{a_1} e^{a_2 x_i}$	$\ln \hat{y}_i = b_0 + b_1 \ln x_i + b_2 x_i$	$a_0 = e^{b_0}, a_1 = b_1, a_2 = b_2$
parabola logarytmiczna	$\hat{y}_i = a_0 x_i^{a_1 + a_2 \ln x_i}$	$\ln \hat{y}_i = b_0 + b_1 \ln x_i + b_2 \ln^2 x_i$	$a_0 = e^{b_0}, a_1 = b_1, a_2 = b_2$

### Metody numeryczne estymacji modeli nieliniowych

Do obliczania ocen parametrów modeli nieliniowych wykorzystuje się często iteracyjne metody numeryczne. Nieliniowa metoda najmniejszych kwadratów polega na wyznaczeniu takich oszacowań parametrów, aby funkcja

$$Q(\theta) = \sum_{i=1}^n [y_i - f(x_{1i}, x_{2i}, \dots, x_{ki}, \theta_1, \theta_2, \dots, \theta_p)]^2$$

osiągała minimum. Tu stosuje się różne procedury iteracyjne; najpopularniejsze z nich to:

- metoda Gaussa-Newtona,
- metoda najszybszego spadku,
- metoda Marquardta.

Metoda Gaussa-Newtona polega na zastąpieniu w kolejnych krokach funkcji  $f(x_{1i}, x_{2i}, \dots, x_{ki}, \theta)$  jej liniową aproksymantą, będącą dwoma pierwszymi składnikami rozwinięcia funkcji w szereg Taylora. Do estymacji parametrów  $\theta_1, \theta_2, \dots, \theta_p$  stosuje się klasyczną metodę najmniejszych kwadratów. W kroku  $(l+1)$  zamiast  $\theta_1^{(l)}, \theta_2^{(l)}, \dots, \theta_p^{(l)}$  umieszcza się oszacowane wartości parametrów:  $\theta_1^{(l+1)}, \theta_2^{(l+1)}, \dots, \theta_p^{(l+1)}$ . Proces iteracyjny trwa aż do osiągnięcia  $|\{\theta_j^{(l+1)} - \theta_j^{(l)}\} / \theta_j^{(l)}| < \delta$ , gdzie  $\delta$  jest z góry ustaloną małą liczbą dodatnią.

W metodzie najszybszego spadku realizuje się proces iteracyjny posuwając się od punktu początkowego  $\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_p^{(0)}$  wzdłuż wektora  $\left[ -\frac{\partial Q(\theta)}{\partial \theta_1}, -\frac{\partial Q(\theta)}{\partial \theta_2}, \dots, -\frac{\partial Q(\theta)}{\partial \theta_p} \right]$ , wektora określającego kierunek najszybszego spadku funkcji  $Q(\theta)$ .

Metoda Marquardta łączy cechy metod linearyzacji Gaussa-Newtona i najszybszego spadku.

## VIII. EKONOMETRYCZNA ANALIZA PRODUKCJI I POPYTU

1. Ekonometryczna analiza produkcji
2. Ekonometryczna analiza popytu

### 1. EKONOMETRYCZNA ANALIZA PRODUKCJI

Podstawowym narzędziem ekonometrycznej analizy procesu produkcyjnego jest funkcja produkcji postaci:  $y = f(x_1, x_2, \dots, x_k, \varepsilon)$ , gdzie  $y$  - wielkość produkcji,  $x_1, x_2, \dots, x_k$  - wielkości czynników produkcji, a  $\varepsilon$  - składnik losowy.

Funkcja produkcji służy do badania zależności między wielkością produkcji, a wielkościami czynników produkcji.

Funkcja produkcji może przyjmować różne postacie analityczne. Zwykle jest funkcją nieliniową, aczkolwiek stosuje się czasem i funkcję liniową, pomimo, że nie spełnia ona pożądanych własności.

Ograniczmy się wstępnie do dwóch czynników:  $L$  – nakładów pracy oraz  $K$  – kapitału. Dla uproszczenia zapisu przyjmijmy:  $Y = f(K, L)$ ,  $Y > 0, K > 0, L > 0$ .

Niech funkcja  $f$  jest dwukrotnie różniczkowalna oraz wykres jej izokwanty będzie wypukły. Można sformułować następujące **założenia, które powinna spełniać funkcja produkcji**:

1. Produkcyjność krańcowa czynnika produkcji jest dodatnia, czyli pochodne cząstkowe pierwszego rzędu względem czynników produkcji muszą być dodatnie:  $f_K > 0, f_L > 0$ . Oznacza to, że zwiększenie zużycia danego czynnika powoduje wzrost produkcji.
2. Produkcyjność krańcowa jest malejąca względem nakładów danego czynnika, czyli pochodne cząstkowe spełniają:  $f_{KK} < 0, f_{LL} < 0$ . Oznacza to, że krańcowe przyrosty maleją w miarę wzrostu danego nakładu, przy nie zmieniającym się drugim nakładzie.
3. Krańcowa produkcyjność jednego czynnika wzrasta przy wzroście drugiego czynnika, czyli pochodne mieszane  $f_{KL} = f_{LK} > 0$ .
4. Funkcja  $f$  jest jednorodna, czyli:  $f(aK, aL) = a^r f(K, L)$ , dla  $a > 0$ . Jednorodność funkcji produkcji określa, jak zareaguje produkcja na zwiększenie wielkości nakładów. Dla  $r=1$  jest to wzrost proporcjonalny, dla  $r > 1$  ma miejsce rosnąca skala produkcji, a dla  $r < 1$  malejąca.
5. Czynniki produkcji są wzajemnie zastępowalne (zachodzi substytucja czynników produkcji).

Do opisu zależności pomiędzy produkcją a czynnikami wytwórczymi najczęściej stosowane są trzy postacie modeli ekonometrycznych:

- **model liniowy:**  $\hat{y}_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + \dots + b_k x_{ki}$
- **model wielomianowy stopnia drugiego:**  $\hat{y}_i = b_0 + b_1 x_{1i} + b_2 x_{1i}^2$
- **model potęgowy:**  $\hat{y}_i = b_0 x_{1i}^{b_1} x_{2i}^{b_2} \dots x_{ki}^{b_k}$ ,  $b_j > 0, j = 0, 1, \dots, k$

Funkcją o wielu zaletach, powodujących częste jej stosowanie do analizy współzależności pomiędzy wielkością produkcji a nakładami, jest potęgowa **funkcja Cobba-Douglasa**.

Postać dwuczynnikowej funkcji produkcji Cobba-Douglasa:

$$y = \beta_0 x_1^{\beta_1} x_2^{\beta_2} e^\varepsilon, \quad \beta_j > 0, j = 0, 1, 2.$$

Funkcja Cobba-Douglasa jest nieliniowym modelem ekonometrycznym, którego parametry możemy oszacować na drodze linearyzacji lub wykorzystując iteracyjne metody estymacji nieliniowej.

## Badanie elastyczności produkcji

Współczynnik elastyczności określa stosunek względnego przyrostu zmiennej objaśnianej do względnego przyrostu zmiennej objaśniającej, przy niezmienionym poziomie pozostałych nakładów.

Wzór na **elastyczność łukową**:

$$E_{y/x} = \frac{\Delta y}{y} : \frac{\Delta x}{x} = \frac{\Delta y}{\Delta x} \cdot \frac{x}{y}.$$

W przypadku znanej postaci analitycznej funkcji produkcji **elastyczność punktową** produkcji  $y$  względem nakładu czynnika  $x_i$  oblicza się według wzoru:

$$E_{\hat{y}/x_i} = \frac{\partial f(x_1, x_2, \dots, x_k)}{\partial x_i} \cdot \frac{x_i}{\hat{y}}.$$

Dla dwuczynnikowej funkcji Cobba – Douglasa  $\hat{y}_i = b_0 x_i^{b_1} x_2^{b_2}$ :

$$E_{\hat{y}/x_1} = b_0 \cdot b_1 x_1^{b_1-1} \cdot x_2^{b_2} \cdot \frac{x_1}{b_0 \cdot x_1^{b_1} \cdot x_2^{b_2}} = b_1 \quad \text{oraz} \quad E_{\hat{y}/x_2} = b_0 \cdot x_1^{b_1} \cdot b_2 x_2^{b_2-1} \cdot \frac{x_2}{b_0 \cdot x_1^{b_1} \cdot x_2^{b_2}} = b_2$$

**Potęgową funkcją produkcji** jest funkcją o **stałej elastyczności produkcji**. Elastyczności względem poszczególnych nakładów czynników produkcji są równe ocenom parametrów modelu ekonometrycznego  $b_i$ , dla  $i = 1, 2, \dots, k$ .

### Przykład 10.

Z oszacowanej funkcji produkcji Cobba - Douglasa  $\hat{y} = 12x_1^{0,45} x_2^{0,67}$  wynika, że elastyczność produkcji względem pierwszego czynnika jest równa 0,45. Zwiększenie wykorzystania tego czynnika o 1% spowoduje przeciętny wzrost produkcji o 0,45%, przy założeniu, że drugi czynnik pozostanie na stałym poziomie.

Suma wartości współczynników elastyczności względem wszystkich czynników produkcji nazywana jest **efektem skali produkcji**:

$$ESP = \sum_{j=1}^k E_{\hat{y}/x_j}.$$

$ESP$  określa procentowy wzrost produkcji, jeżeli czynniki produkcji wzrosną o 1%.

- Jeżeli  $ESP = 1$ , to ma miejsce stała wydajność czynników produkcji (stałej korzyści skali), czyli produkcja wzrasta w tym samym tempie co nakłady czynników produkcji.
- Jeżeli  $ESP > 1$ , to ma miejsce rosnąca wydajność czynników produkcji (rosnące korzyści skali), czyli tempo przyrostu produkcji jest wyższe niż czynników produkcji.
- Jeżeli  $ESP < 1$ , to ma miejsce malejąca wydajność czynników produkcji (malejące przychody skali), czyli produkcja przyrasta w tempie wolniejszym niż nakłady.

Z przypadkami stałej, malejącej i rosnącej skali produkcji związane jest pojęcie jednorodności funkcji. Potęgowa funkcja produkcji Cobba-Douglasa jest jednorodna, a stopień jednorodności tej funkcji określa wysokość efektu skali produkcji.

Warto zauważyć, że w dwuczynnikowym **liniowym modelu produkcji elastyczność produkcji** względem nakładów  $x_1$  i  $x_2$  **nie jest stała**, ale zależy od poziomu wykorzystania czynników:

$$EP_{x_1} \Big|_{\substack{x_1=x_1^0 \\ x_2=x_2^0}} = \frac{b_1 x_1^0}{b_0 + b_1 x_1^0 + b_2 x_2^0}, \quad EP_{x_2} \Big|_{\substack{x_1=x_1^0 \\ x_2=x_2^0}} = \frac{b_2 x_2^0}{b_0 + b_1 x_1^0 + b_2 x_2^0}$$

**Izokwanta produkcji** to linia jednakowej produkcji, przedstawia ona wszystkie możliwe kombinacje czynników produkcji, które pozwalają osiągnąć tą samą wielkość produkcji.

### Badanie przyrostów krańcowych produkcji

Produkcyjność krańcowa czynnika mówi, o ile przeciętnie zmieni się wielkość produkcji na jednostkę zużycia czynnika produkcji, przy nie zmienionych nakładach pozostałych czynników. Produkcyjność krańcowa to odpowiednia pochodna cząstkowa funkcji produkcji względem  $k$ -tego czynnika:

$$PK_{x_k} = \frac{\partial \hat{y}}{\partial x_k}$$

Dla dwuczynnikowego **liniowego modelu produkcji** przyrosty krańcowe zmiennej objaśnianej względem  $j$ -tej zmiennej objaśniającej wynoszą:

$$PK_{x_1} = \frac{\partial \hat{y}}{\partial x_1^0} = b_1, \quad PK_{x_2} = \frac{\partial \hat{y}}{\partial x_2^0} = b_2$$

i są wielkościami **stałymi** dla wszystkich wartości  $x_j$ .

Przyrosty krańcowe dla dwuczynnikowej **funkcji potęgowej** wynoszą:

$$PK_{x_1} \Big|_{\substack{x_1=x_1^0 \\ x_2=x_2^0}} = b_0 \cdot b_1 (x_1^0)^{b_1-1} \cdot (x_2^0)^{b_2}, \quad PK_{x_2} \Big|_{\substack{x_1=x_1^0 \\ x_2=x_2^0}} = b_0 \cdot b_2 (x_1^0)^{b_1} \cdot (x_2^0)^{b_2-1}$$

i są **różne** w zależności od wartości nakładów.

### Badanie substytucji czynników produkcji

Substytucja czynników produkcji polega na możliwości zastępowania zużycia jednego z czynników innymi czynnikami bez zmiany wielkości produkcji.

**Krańcowa stopa substytucji czynnika  $x_1$  względem czynnika  $x_2$ :**

$$KSS_{x_1x_2} = \frac{dx_2}{dx_1} = -\frac{PK_{x_1}}{PK_{x_2}}$$

Krańcowa stopa substytucji czynnika  $x_1$  względem czynnika produkcji  $x_2$  (czyli zastępowanie czynnika  $x_1$  przez czynnik  $x_2$ ) określa, ile czynnika  $x_2$  musi być wprowadzone w miejsce wycofanej jednostki czynnika  $x_1$ , tak aby produkcja pozostała na tym samym poziomie.

Dla liniowej funkcji produkcji: 
$$KSS_{x_1x_2} = \frac{dx_2}{dx_1} = -\frac{PK_{x_1}}{PK_{x_2}} = -\frac{b_1}{b_2}$$

Dla potęgowej funkcji produkcji: 
$$KSS_{x_1x_2} = \frac{dx_2}{dx_1} = -\frac{PK_{x_1}}{PK_{x_2}} = -\frac{b_0 \cdot b_1 x_1^{b_1-1} \cdot x_2^{b_2}}{b_0 \cdot x_1^{b_1} \cdot b_2 x_2^{b_2-1}} = -\frac{b_1}{b_2} \cdot \frac{x_2}{x_1}$$

### Przykład 11.

$KSS_{LK} = -0,2$  oznacza, że przy spadku nakładu pracy  $L$  o jednostkę, należy zwiększyć nakład kapitału  $K$  o 0,2 jednostki, tak aby poziom produkcji końcowej pozostał na tym samym poziomie

### Dynamiczne funkcje produkcji

Do badań mikroekonomicznych w krótkim okresie można wykorzystać statyczną funkcję produkcji. Jednak gdy wielkość produkcji zależy od czasu, to funkcja produkcji winna mieć charakter dynamiczny. Dynamiczna postać dwuczynnikowej funkcji produkcji:

$$y_t = \gamma x_{1t}^\alpha x_{2t}^\beta e^{\delta t + \varepsilon_t}$$



Parametr  $\delta$  jest miarą postępu techniczno-organizacyjnego. Jeżeli  $\delta > 0$ , to wielkość produkcji zależy od efektów tego postępu. Wielkość wzrostu produkcji wynosi wtedy średnio  $(e^\delta - 1) \cdot 100\%$  w skali roku.

## 2. EKONOMETRYCZNA ANALIZA POPYTU

Badania popytu konsumpcyjnego zapoczątkował w XIX w. Engel. Udowodnił on, że w miarę wzrostu dochodów ludności następuje spadek udziału wydatków na cele pierwszej potrzeby (np. żywność) i wzrost udziału wydatków na cele wyższego rzędu (np. dobra luksusowe).

Zdefiniujmy **współczynnik elastyczności popytu** względem  $i$ -tego czynnika:

$$E_{\hat{y}/x_i} = \frac{\partial \hat{y}}{\partial x_i} \cdot \frac{x_i}{\hat{y}}$$

Określa on, o ile średnio wzrasta lub maleje popyt, gdy czynnik  $x_i$  wzrasta o 1%, przy założeniu stałości pozostałych czynników.

**Dochodowa elastyczność popytu** - siła reakcji popytu na zmianę dochodu.

**Cenowa elastyczność popytu** - siła reakcji popytu na zmianę ceny.

Zazwyczaj współczynniki dochodowej elastyczności popytu są dodatnie, a cenowej ujemne.

Jeśli  $\left| E_{\hat{y}/x_i} \right| > 1$ , to popyt jest doskonale elastyczny,  $\left| E_{\hat{y}/x_i} \right| = 1$  - popyt proporcjonalny,

$0 < \left| E_{\hat{y}/x_i} \right| < 1$  - popyt mało elastyczny, a  $\left| E_{\hat{y}/x_i} \right| = 0$  określa popyt sztywny.

Do badań popytu konsumpcyjnego najczęściej stosuje się następujące modele:

- **Model liniowy:**  $\hat{y}_i = b_0 + b_1 x_{1i} + \dots + b_k x_{ki}$
- **Model potęgowy:**  $\hat{y}_i = b_0 x_{1i}^{b_1} \dots x_{ki}^{b_k}$
- **Model Workinga:**  $\hat{y}_i = \exp(b_0 + \frac{b_1}{x_i})$

Stosowany w badaniach wydatków na dane dobro w zależności od wydatków ogółem.

- **Model Törnquista I:**  $\hat{y}_i = \frac{ax_i}{b + x_i}$ ,  $a > 0, b > 0$

Stosowany do badania zależności pomiędzy dochodem a wydatkami na dobra pierwszej potrzeby.

- **Model Törnquista II:**  $\hat{y}_i = \frac{a(x_i - c)}{x_i + b}$ ,  $a > 0, b > 0, c > 0$

Stosowany do opisu zależności pomiędzy dochodem a wydatkami na dobra wyższego rzędu. Wielkość parametru  $a$  wskazuje na poziom nasycenia, czyli poziom do którego wydatki rosną, lecz którego nigdy nie przekroczyć. Parametr  $c$  wskazuje na poziom dochodów, przy którym pojawiają się wydatki na badane dobro.

- **Model Törnquista III:**  $\hat{y}_i = \frac{ax_i(x_i - c)}{x_i + b}$ ,  $a > 0, b > 0, c > 0$

Stosowany do opisu zależności pomiędzy dochodem a wydatkami na dobra luksusowe. Parametr  $c$  wskazuje poziom dochodu, od którego pojawia się popyt na dobra luksusowe.

Estymacja parametrów modeli Workinga i Törnquista:

- etap I: sprowadzamy modele do postaci liniowej i oszacujemy ich parametry MNK,
- etap II: uzyskane oceny parametrów w etapie I stają się wartościami początkowymi w nieliniowej MNK.

Przykładowo model Törnquista I sprowadza się do postaci liniowej przekształcając równanie modelowe do postaci:  $\frac{1}{y_i} = \frac{x_i + b}{ax_i} = \frac{1}{a} + \frac{b}{ax}$ . Po podstawieniach  $u_i = \frac{1}{y_i}$ ,  $\beta_0 = \frac{1}{a}$ ,  $\beta_1 = \frac{b}{a}$ ,

$z_i = \frac{1}{x_i}$  otrzymujemy model liniowy  $u_i = \beta_0 + \beta_1 z_i + \varepsilon_i$ . Szacujemy go KMNK. Uzyskane oceny będą parametrami startowymi w iteracyjnej metodzie estymacji.

Podamy teraz, jak obliczyć dochodową elastyczność popytu na podstawie modelu:

- Workinga: 
$$E_{\hat{y}/x} = \frac{d\hat{y}}{dx} \cdot \frac{x}{\hat{y}} = \exp\left(a + b \cdot \frac{1}{x}\right) \cdot \left(\frac{-b}{x^2}\right) \cdot \frac{x}{\exp\left(a + b \cdot \frac{1}{x}\right)} = \frac{-b}{x}$$
- Törnquista I: 
$$E_{\hat{y}/x} = \frac{d\hat{y}}{dx} \cdot \frac{x}{\hat{y}} = \frac{a(x+b) - 1 \cdot ax}{(x+b)^2} \cdot \frac{x(x+b)}{ax} = \frac{b}{x+b}$$
- Törnquista II: 
$$E_{\hat{y}/x} = \frac{d\hat{y}}{dx} \cdot \frac{x}{\hat{y}} = \frac{a(x+b) - 1 \cdot a(x-c)}{(x+b)^2} \cdot \frac{x(x+b)}{a(x-c)} = \frac{x(b-c)}{(x+b)(x-c)}$$
- Törnquista III: 
$$E_{\hat{y}/x} = \frac{d\hat{y}}{dx} \cdot \frac{x}{\hat{y}} = \left(\frac{ax^2 - acx}{x+b}\right)' \cdot \frac{x(x+b)}{ax(x-c)} = \frac{x^2 + 2bx - bc}{(x+b)(x-c)} = \frac{x}{x-c} + \frac{b}{x+b}$$

### Przykład 12.

Posługując się modelem Törnquista I wyznaczono współczynnik elastyczności wydatków na konsumpcję względem dochodów otrzymując  $E_{\hat{y}/x} = 0,73$ . Oznacza to, że wzrost dochodów o 1%, powoduje wzrost średniego poziomu wydatków na konsumpcję o około 0,73%.



## IX. WIELORÓWNANIOWE MODELE EKONOMETRYCZNE

1. Klasyfikacja zmiennych
2. Postać strukturalna i zredukowana modelu
3. Klasyfikacja modeli wielorównaniowych
4. Estymacja parametrów modeli wielorównaniowych

### 1. KLASYFIKACJA ZMIENNYCH

W wypadku złożonych zjawisk ekonomicznych model jednorównaniowy tłumaczący kształtowanie się jednej zmiennej objaśnianej może być niewystarczający. W takim wypadku konstruujemy model wielorównaniowy - zespół równań. Każde z równań wyjaśnia kształtowanie się jednej zmiennej objaśnianej, a zmienna objaśniana w jednym równaniu, może być objaśniającą w innym równaniu.

**Model wielorównaniowy** jest modelem zawierającym co najmniej 2 równania. Przynajmniej jedno równanie w modelu musi być stochastyczne.

#### Przykład 13.

Następujący model makroekonomiczny jest przykładem wielorównaniowego modelu ekonometrycznego:

$$C_t = \alpha_0 + \alpha_1 Y_t + \alpha_2 C_{t-1} + \varepsilon_{1t}$$

$$I_t = \beta_0 + \beta_1 R_t + \beta_2 (Y_t - Y_{t-1}) + \varepsilon_{2t}$$

$$Y_t = C_t + I_t + G_t$$

gdzie:

$C_t$  – konsumpcja ogółem, np. realne wydatki konsumpcyjne w danej gospodarce w roku  $t$ ,

$Y_t$  – globalny dochód, np. produkt krajowy brutto, wartość realna w roku  $t$ ,

$I_t$  – inwestycje, np. nakłady brutto na środki trwałe, w cenach stałych w roku  $t$ ,

$R_t$  – stopa procentowa, np. stopa lombardowa NBP w roku  $t$ ,

$G_t$  – wydatki rządowe w roku  $t$ ,

$C_{t-1}$  oraz  $Y_{t-1}$  – konsumpcja i globalny dochód (zagregowany popyt) z roku poprzedniego,

$\varepsilon_{1t}, \varepsilon_{2t}$  – składniki losowe.

Inny model - model rynku w równowadze - składa się z 3 równań:

$$Q_t^d = \alpha_0 + \alpha_1 P_t + \varepsilon_{1t}$$

$$Q_t^s = \beta_0 + \beta_1 P_t + \varepsilon_{2t}$$

$$Q_t^d = Q_t^s$$

gdzie:

$Q_t^d, Q_t^s$  – wielkość popytu i podaży w roku  $t$ ,

$P_t$  – cena określonego dobra,

$\varepsilon_{1t}, \varepsilon_{2t}$  – składniki losowe.

W modelach wielorównaniowych występują zmienne **endogeniczne** i **egzogeniczne**. Zmienne endogeniczne to te, które są objaśniane w jakimś równaniu modelu. Zmienne egzogeniczne określone są poza modelem. W dynamicznym modelu wielorównaniowym mogą występować zmienne z różnych okresów. Endogeniczne i egzogeniczne zmienne

mogą być **nieopóźnione**, czyli z okresu bieżącego, jak i **opóźnione**, czyli z okresów wcześniejszych.

Zmienne endogeniczne nieopóźnione nazywa się zmiennymi **łącznie współzależnymi**. Opóźnione zmienne endogeniczne i zmienne egzogeniczne nazywamy zmiennymi **z góry ustalonymi**.

Zmienne	Nieopóźnione	Opóźnione
Endogeniczne	Łącznie współzależne	
Egzogeniczne		Z góry ustalone

## 2. POSTAĆ STRUKTURALNA I ZREDUKOWANA MODELU

Istnieją dwa sposoby zapisu modeli wielorównaniowych. Zapis w postaci strukturalnej i zapis w postaci zredukowanej.

**Postać strukturalna** oddaje rzeczywiste relacje między zmiennymi. Każde równanie w tej postaci objaśnia jedną zmienną pozostałymi zmiennymi.

Oznaczmy:

- $y_{it}$  – obserwacja  $i$ -tej zmiennej endogenicznej w okresie  $t$ ,  $i = 1, 2, \dots, m$ ,
- $z_{jt}$  – obserwacja  $j$ -tej zmiennej z góry ustalonej w okresie  $t$ ,  $j = 1, 2, \dots, k$ ,
- $\beta_{il}$  – parametr przy  $l$ -tej nieopóźnionej zmiennej endogenicznej w  $i$ -tym równaniu,  $l, i = 1, 2, \dots, m$ ,
- $\gamma_{ij}$  – parametr przy  $j$ -tej zmiennej z góry ustalonej w  $i$ -tym równaniu,  $i = 1, 2, \dots, m$ ,  $j = 1, 2, \dots, k$ .

Jeśli w pewnym równaniu modelowym miałyby wystąpić stała, to wprowadzamy fikcyjną zmienną stałe równą jeden.

Rozważmy model składający się z  $m$  równań:

$$y_{1t} = \sum_{l=2}^m \beta_{1l} y_{lt} + \sum_{j=1}^k \gamma_{1j} z_{jt} + \varepsilon_{1t}$$

$$y_{2t} = \sum_{\substack{l=1 \\ l \neq 2}}^m \beta_{2l} y_{lt} + \sum_{j=1}^k \gamma_{2j} z_{jt} + \varepsilon_{2t}$$

.....

$$y_{mt} = \sum_{l=1}^{m-1} \beta_{ml} y_{lt} + \sum_{j=1}^k \gamma_{mj} z_{jt} + \varepsilon_{mt}$$

Przenosząc zmienne współzależne i z góry ustalone na jedną stronę dostajemy:

$$y_{1t} - \sum_{l=2}^m \beta_{1l} y_{lt} - \sum_{j=1}^k \gamma_{1j} z_{jt} = \varepsilon_{1t}$$

$$y_{2t} - \sum_{\substack{l=1 \\ l \neq 2}}^m \beta_{2l} y_{lt} - \sum_{j=1}^k \gamma_{2j} z_{jt} = \varepsilon_{2t}$$

.....

$$y_{mt} - \sum_{l=1}^{m-1} \beta_{ml} y_{lt} - \sum_{j=1}^k \gamma_{mj} z_{jt} = \varepsilon_{mt}$$



W zapisie macierzowym powyższa postać strukturalna jest następująca:

$$\mathbf{B}\mathbf{Y}_t + \mathbf{\Gamma}\mathbf{Z}_t = \boldsymbol{\varepsilon}_t, \quad t = 1, 2, \dots, n,$$

gdzie

$$\mathbf{B} = \begin{bmatrix} 1 & -\beta_{12} & \dots & -\beta_{1m} \\ -\beta_{21} & 1 & \dots & -\beta_{2m} \\ \dots & \dots & \dots & \dots \\ -\beta_{m1} & -\beta_{m2} & \dots & 1 \end{bmatrix}, \quad \mathbf{\Gamma} = \begin{bmatrix} -\gamma_{11} & -\gamma_{12} & \dots & -\gamma_{1k} \\ -\gamma_{21} & -\gamma_{22} & \dots & -\gamma_{2k} \\ \dots & \dots & \dots & \dots \\ -\gamma_{m1} & -\gamma_{m2} & \dots & -\gamma_{mk} \end{bmatrix},$$

$$\mathbf{Y}_t = \begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{mt} \end{bmatrix}, \quad \mathbf{Z}_t = \begin{bmatrix} z_{1t} \\ z_{2t} \\ \vdots \\ z_{kt} \end{bmatrix}, \quad \boldsymbol{\varepsilon}_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{mt} \end{bmatrix}.$$

W **postaci zredukowanej** modelu w każdym równaniu występuje tylko jedna zmienna endogeniczna nieopóźniona, ta która jest wyjaśniana równaniem, wszystkie pozostałe zmienne są z góry ustalone. Zakładamy istnienie macierzy odwrotnej  $\mathbf{B}^{-1}$ . Dokonajmy przekształcenia równania  $\mathbf{B}\mathbf{Y}_t + \mathbf{\Gamma}\mathbf{Z}_t = \boldsymbol{\varepsilon}_t$ :

$$\mathbf{B}^{-1}\mathbf{B}\mathbf{Y}_t + \mathbf{B}^{-1}\mathbf{\Gamma}\mathbf{Z}_t = \mathbf{B}^{-1}\boldsymbol{\varepsilon}_t.$$

$$\mathbf{Y}_t = -\mathbf{B}^{-1}\mathbf{\Gamma}\mathbf{Z}_t + \mathbf{B}^{-1}\boldsymbol{\varepsilon}_t.$$

Oznaczamy  $\mathbf{\Pi} = -\mathbf{B}^{-1}\mathbf{\Gamma}$ ,  $\boldsymbol{\eta}_t = \mathbf{B}^{-1}\boldsymbol{\varepsilon}_t$  i otrzymujemy **postać zredukowaną modelu**:

$$\mathbf{Y}_t = \mathbf{\Pi}\mathbf{Z}_t + \boldsymbol{\eta}_t.$$

Inaczej:

$$y_{1t} = \pi_{11}z_{1t} + \pi_{12}z_{2t} + \dots + \pi_{1k}z_{kt} + \eta_{1t}$$

$$y_{2t} = \pi_{21}z_{1t} + \pi_{22}z_{2t} + \dots + \pi_{2k}z_{kt} + \eta_{2t}$$

.....

$$y_{mt} = \pi_{m1}z_{1t} + \pi_{m2}z_{2t} + \dots + \pi_{mk}z_{kt} + \eta_{mt}$$

Postać zredukowana jest wygodna do szacowania parametrów modelu.

### 3. KLASYFIKACJA MODELI WIELORÓWNANIOWYCH

Modele wielorównaniowe dzielimy **na trzy rozłączne klasy** ze względu na powiązania występujące między nieopóźnionymi zmiennymi endogenicznymi:

- modele proste,
- modele rekurencyjne,
- modele o równaniach współzależnych.

Model jest **modelem prostym**, jeśli zmienne endogeniczne nieopóźnione nie są ze sobą wzajemnie powiązane.

Model jest **modelem rekurencyjnym**, jeśli powstaje łańcuch zmiennych endogenicznych nieopóźnionych (powiązania pomiędzy zmiennymi modelu są jednokierunkowe). W równaniach objaśniających zmienną  $Y_k$  pojawiają się zmienne  $Y_j$  dla  $j < k$ , nie pojawiają się zaś zmienne  $Y_j$  dla  $j > k$ .

Model jest **modelem o równaniach współzależnych**, jeśli nie jest ani modelem prostym, ani rekurencyjnym. W tym modelu pojawiają się zmienne nieopóźnione endogeniczne jednocześnie od siebie zależne (pomiędzy zmiennymi endogenicznymi istnieje co najmniej jedno sprzężenie zwrotne).

Jeżeli macierz **B** współczynników przy nieopóźnionych zmiennych endogenicznych jest:

- diagonalna, to rozważany model jest modelem prostym,
- trójkątna lub może być sprowadzona do trójkątnej poprzez przenieśnięcie zmiennych i równań w modelu, to rozważany model jest modelem rekurencyjnym,
- ani diagonalna, ani trójkątna, to model jest modelem o równaniach współzależnych.

#### 4. ESTYMACJA PARAMETRÓW MODELI WIELORÓWNANIOWYCH

**Modele proste**, charakteryzujące się brakiem powiązań między zmiennymi endogenicznymi nieopóźnionymi występującymi w poszczególnych równaniach modelu, możemy traktować tak, jakby składały się z osobnych modeli jednorównaniowych. Wtedy osobno każde równanie szacujemy klasyczną metodą najmniejszych kwadratów.

W przypadku **modelu rekurencyjnego** przenieśnięmy równania tak, aby macierz parametrów przy nieopóźnionych zmiennych endogenicznych była macierzą trójkątną. Szacujemy następnie parametry pierwszego równania, w którym w charakterze zmiennych objaśniających występują jedynie zmienne z góry ustalone. Na podstawie oszacowanych parametrów pierwszego równania wyznaczamy wartości teoretyczne nieopóźnionej zmiennej endogenicznej. W kolejnych równaniach, gdzie w charakterze zmiennych objaśniających występują nieopóźnione zmienne endogeniczne, przy estymacji parametrów zamiast wartości empirycznych uwzględniamy wartości teoretyczne tych zmiennych.

Gdy model jest **modelem o równaniach łącznie współzależnych**, sprowadzamy model do postaci zredukowanej. W postaci zredukowanej badany model jest modelem prostym. Jeśli składniki losowe z różnych równań są niezależne stosujemy do każdego z nich MNK.

Nie zawsze jednak na podstawie parametrów postaci zredukowanej modelu można określić parametry postaci strukturalnej. Problem ten znany jest pod nazwą **problemu identyfikacji**.

Model jest **identyfikowalny** jeśli na podstawie parametrów postaci zredukowanej można wyznaczyć wszystkie parametry postaci strukturalnej.

Dane równanie strukturalne jest identyfikowalne, jeśli wszystkie jego parametry można wyznaczyć na podstawie znajomości postaci zredukowanej. Model jest identyfikowalny, gdy każde równanie strukturalne jest identyfikowalne.

W problemie identyfikacji możliwe są 3 sytuacje:

1. Nie jest możliwe wyznaczenie wszystkich parametrów równania strukturalnego na podstawie znajomości parametrów postaci zredukowanej modelu – równanie jest **nieidentyfikowalne**.
2. Na podstawie znajomości parametrów postaci zredukowanej modelu można w sposób jednoznaczny określić parametry równania strukturalnego – równanie jest **jednoznacznie identyfikowalne**.
3. Jeśli parametry równania strukturalnego można wyznaczyć na podstawie parametrów postaci zredukowanej w sposób niejednoznaczny, to mówimy o równaniu **niejednoznacznie identyfikowalnym**.

Problem identyfikacji polega na rozstrzygnięciu, czy układ równań  $\mathbf{B}\Pi = -\Gamma$  ma rozwiązanie; jeśli tak, to jakie: jednoznaczne, czy niejednoznaczne?

**Warunek wymiaru:** Warunkiem koniecznym identyfikowalności danego równania jest, by liczba zmiennych z góry ustalonych niewystępujących w tym równaniu była większa bądź równa liczbie nieopóźnionych zmiennych endogenicznych występujących w tym równaniu pomniejszonej o jeden. Jeśli okaże się równa, to równanie może być jednoznacznie identyfikowalne, natomiast jeśli większa to może być niejednoznacznie identyfikowalne.



**Warunek rzędu:** Warunkiem koniecznym i dostatecznym na to, aby  $i$ -te równanie modelu składającego się z  $m$  równań było identyfikowalne, jest by macierz utworzona z parametrów znajdujących się przy zmiennych, które są w modelu, ale nie występują w  $i$ -tym równaniu była rzędu  $m-1$ .

Identyfikowalność modelu należy sprawdzić przed przystąpieniem do estymacji jego parametrów. Jeśli okaże się, że pewne równanie modelu nie jest identyfikowalne, wówczas jego parametrów nie da się oszacować. Również nie każda metoda estymacji może być zastosowana w przypadku niejednoznacznej identyfikowalności. W przypadku modeli identyfikowalnych albo jednoznacznie, albo niejednoznacznie, metodą szacowania parametrów jest **podwójna metoda najmniejszych kwadratów (2MNK)**. Inną popularną metodą szacowania parametrów modelu jednoznacznie identyfikowalnego jest **pośrednia metoda najmniejszych kwadratów**.