

Ekonometria

Mieczysław Sobczyk

Wydawnictwo C.H. Beck 

METODY ILOŚCIOWE

M. Sobczyk
Ekonometria

Ekonometria

Mieczysław Sobczyk

wydanie 1



WYDAWNICTWO C.H. BECK
WARSZAWA 2012

Wydawca: Dorota Ostrowska-Furmanek
Redakcja merytoryczna: Beata Socha
Recenzent: dr hab. Maria Balcerowicz-Szkutnik, prof. UE
Projekt okładki i stron tytułowych: Maryna Wiśniewska
Ilustracja na okładce: © Mark Evans/iStockphoto.com

Seria: Metody ilościowe

Złożono programem T_EX



© Wydawnictwo C.H. Beck 2012

Wydawnictwo C.H. Beck Sp. z o.o.
ul. Bonifraterska 17, 00-203 Warszawa

Skład i łamanie: Wydawnictwo C.H. Beck
Druk i oprawa: Elpil, Siedlce

ISBN 978-83-255-4344-0



e-book 978-83-255-4345-7

Spis treści

Wprowadzenie	7
Rozdział 1. Modelowanie ekonometryczne	9
1.1. Istota modelu ekonometrycznego i jego elementy składowe	9
1.2. Etapy budowy modelu ekonometrycznego	11
1.3. Klasyfikacja modeli ekonometrycznych	15
Rozdział 2. Dobór zmiennych objaśniających do modelu ekonometrycznego .	17
2.1. Współczynnik zmienności jako kryterium doboru zmiennych	17
2.2. Metoda analizy współczynników korelacji jako podstawa doboru zmiennych	20
2.3. Wykorzystanie współczynników korelacji wielorakiej w celu doboru zmiennych	22
2.4. Dobór zmiennych objaśniających metodą Hellwiga	24
Zadania	27
Rozdział 3. Estymacja parametrów liniowego modelu ekonometrycznego . .	33
3.1. Pojęcie estymacji i estymatora. Założenia klasycznej metody najmniejszych kwadratów (KMNK)	33
3.2. Estymacja parametrów modelu z jedną zmienną objaśniającą	36
3.3. Estymacja parametrów modelu liniowego z wieloma zmiennymi objaśniającymi	41
Zadania	48
Rozdział 4. Weryfikacja jednorównaniowego liniowego modelu ekonometrycznego	53
4.1. Istotność parametrów strukturalnych	53
4.2. Badanie normalności składnika losowego	58
4.3. Badanie autokorelacji składnika losowego	62
4.4. Badanie jednorodności wariancji składnika losowego	66
Zadania	67
Rozdział 5. Wykorzystanie jednorównaniowych modeli ekonometrycznych do prognozowania	71
5.1. Pojęcie, funkcje i klasyfikacja prognoz	71
5.2. Zasady i metody prognozowania	73
5.3. Budowa prognoz punktowych i przedziałowych	75
5.4. Mierniki dokładności predykcji	86
Zadania	99

Rozdział 6. Nieliniowe modele ekonometryczne	106
6.1. Typy modeli nieliniowych	106
6.2. Estymacja parametrów modeli nieliniowych	106
6.2.1. Modele liniowe względem parametrów	106
6.2.2. Modele linearyzowane przez logarytmowanie	113
Zadania	122
Odpowiedzi do zadań	124
Rozdział 2	124
Rozdział 3	124
Rozdział 4	125
Rozdział 5	126
Rozdział 6	126
Tablice	128
Bibliografia	139

Wprowadzenie

Zajęcia z ekonometrii są – w mniejszym lub większym zakresie – realizowane we wszystkich uczelniach ekonomicznych. Przekazywane w ramach tego przedmiotu treści są dość zróżnicowane i uzależnione głównie od kierunku studiów. Niniejszy podręcznik zawiera podstawowy kurs ekonometrii i jest przeznaczony dla studentów różnych dyscyplin ekonomicznych. Sądzę, że książka będzie również przydatna tym wszystkim, którzy prowadzą analizy ekonomiczne, a nie dysponują szeroką wiedzą matematyczno-statystyczną. Prezentowane w podręczniku zagadnienia są bowiem podawane od podstaw i nie są nadmiernie skomplikowane pod względem formalnym. Cenną zaletą podręcznika jest prezentacja przykładów, które ilustrują istotę wykorzystywanych procedur oraz pokazują możliwości interpretacyjne otrzymanych wyników. Ponadto, praca zawiera zestaw zadań do samodzielnego rozwiązywania. Kontrolę stopnia opanowania wiedzy ułatwiają zamieszczone odpowiedzi do zadań.

Zawarte w podręczniku treści zostały ujęte w sześciu merytorycznych rozdziałach. Pierwszy z nich wprowadza Czytelnika w podstawowe zagadnienia dotyczące modelowania ekonometrycznego (jego istotę, elementy składowe, etapy budowy i klasyfikacje).

Rozdział drugi poświęcono prezentacji wybranych metod doboru zmiennych objaśniających do modelu ekonometrycznego (w oparciu o współczynniki: zmienności i korelacji wielorakiej, analizę macierzy współczynników korelacji oraz szeroko wykorzystywaną w praktyce metodę Hellwiga).

W rozdziale trzecim zaprezentowano zagadnienia dotyczące wykorzystania metody najmniejszych kwadratów do estymacji parametrów strukturalnych liniowych modeli ekonometrycznych z jedną i wieloma zmiennymi objaśniającymi.

Rozważania zawarte w rozdziale czwartym skoncentrowane są na problematyce weryfikacji jednorównaniowego modelu ekonometrycznego, w tym również funkcji trendu liniowego. Zostały tu przedstawione techniki z zakresu zastosowania testów statystycznych do badania istotności parametrów strukturalnych, badania normalności, autokorelacji i homoskedastyczności składników losowych.

Przedmiotem zainteresowań w rozdziale piątym jest wykorzystanie modeli ekonometrycznych dla celów prognozowania (predykcji) punktowego i przedziałowego. W końcowej części tego rozdziału zwrócono szczególną uwagę na prezentację mierników bezwzględnych i względnych średnich błędów prognoz *ex ante*.

W ostatnim, szóstym, rozdziale przedstawiono problematykę estymacji, weryfikacji i wykorzystania do prognozowania modeli nieliniowych, które uprzednio należy zlinearyzować (np. przez podstawianie bądź logarytmowanie).

W podręczniku zamieszczono także bibliografię, zawierającą pozycje z dziedziny ekonometrii. Pozwoli to na pogłębienie znajomości zagadnień prezentowanych w niniejszej pracy, jak również – w razie konieczności – na poszerzenie wiedzy w zakresie treści pominiętych.

Podręcznik zawiera także tablice statystyczne, które są niezbędne przy wnioskowaniu statystycznym. W ten sposób tworzy on pewną całość (teoria – przykłady – zadania). Wydaje się, że zarówno zawarty w pracy zakres przedmiotowy, jak i zwięzły oraz jasny sposób ujęcia rozpatrywanych zagadnień – nadają jej przyjazny dla czytelników charakter.

Autor składa gorące podziękowania Pani Profesor Marii Balcerowicz-Szcutnik za cenne uwagi, które miały istotny wpływ na ostateczny kształt podręcznika. Słowa wdzięczności kieruję również do Pani Redaktor Beaty Sochy za niezwykle umiejętność dostrzeżenia wszelkich niedociągnięć oraz do córki Eweliny – za trud włożony w techniczne przygotowanie tekstu.

Mieczysław Sobczyk

Rozdział 1. Modelowanie ekonometryczne

1.1. Istota modelu ekonometrycznego i jego elementy składowe

Istotą modelowania ekonometrycznego jest budowa modelu wyjaśniającego mechanizm zmian zachodzących w badanym wycinku rzeczywistości. Przedmiotem naszych zainteresowań będą ilościowe zależności zachodzące między analizowanymi zjawiskami gospodarczymi (procesami ekonomicznymi).

Z formalnego punktu widzenia **model ekonometryczny** jest równaniem (układem równań) opisującym zasadnicze powiązania między określonymi wielkościami ekonomicznymi. W ogólnej postaci jednorodnaniowy model ekonometryczny jest zapisywany następująco:

$$Y = f(X_1, X_2, \dots, X_k, \varepsilon), \quad (1.1)$$

gdzie: Y – zmienna objaśniana (endogeniczna), X_k – zmienne objaśniające (egzogeniczne), f – analityczna postać funkcji zmiennych objaśniających (np. funkcja liniowa, wykładnicza, parabola), ε – składnik losowy (odchylenia losowe) modelu ekonometrycznego.

Jak wynika z relacji (1.1), model ekonometryczny opisuje powiązanie między zmienną objaśnianą (Y) a zmiennymi objaśniającymi (X_1, X_2, \dots, X_k). Zjawiska ekonomiczne często odznaczają się dużą złożonością. Z tego też względu model ekonometryczny ma jedynie przybliżony charakter. Z reguły nie ma bowiem możliwości skonstruowania prostych, dających się łatwo zinterpretować modeli ekonometrycznych całkowicie zgodnych z danymi empirycznymi. Z drugiej strony, modelem ekonometrycznym nie może być dowolna funkcja. Musi być ona dobrze uzasadniona teoretycznie (z punktu widzenia teorii ekonomii i postulatów ekonometrycznych) i empirycznie.

Uwzględnienie w modelu ekonometrycznym odchyłeń losowych (ε) nadaje mu **stochastyczny** charakter. Składnik losowy odzwierciedla wpływ zjawisk ubocznych, a jego wprowadzenie do modelu ekonometrycznego jest uwarunkowane poniższymi przyczynami [Osińska, 2007, s. 15]:

- przyjęciem niewłaściwej postaci analitycznej funkcji,
- brakiem możliwości uwzględnienia w modelu wszystkich zmiennych objaśniających opisujących kształtowanie się zmiennej objaśnianej,

- błędami wynikającymi z niedokładności pomiaru zmiennych,
- losowością postępowania podmiotów ekonomicznych, a zwłaszcza zachowań ludzkich. Przejawia się to tym, że ten sam konsument w obliczu tak samo sformułowanego dylematu wyboru każdorazowo może podjąć nieco inną decyzję.

Stochastyczny charakter modelu oznacza, że prawidłowości rozwoju badanej zmiennej uwidaczniają się w dużej liczbie obserwacji (prawo wielkich liczb). Im model ekonometryczny lepiej odzwierciedla badaną rzeczywistość, tym odchylenia rzeczywistych wartości zmiennej objaśnianej od jej wartości wyznaczonych z modelu (przy przyjętym zbiorze zmiennych objaśniających) są mniejsze.

Zakładając **liniową zależność** zmiennej objaśnianej Y od zmiennych objaśniających (X_1, X_2, \dots, X_k) , model (1.1) zapisujemy następująco:

$$Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_k X_k + \varepsilon. \quad (1.2)$$

W relacji (1.2) można wyróżnić dwie części: deterministyczną i stochastyczną. Część **deterministyczną** tworzą $\alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_k X_k$, natomiast część stochastyczna związana jest ze składnikiem losowym, czyli zakłócającym (ε).

Po odrzuceniu odchyłeń losowych ε otrzymujemy równanie o postaci:

$$\hat{Y} = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_k X_k. \quad (1.3)$$

W równaniu tym symbol \hat{Y} oznacza oczekiwaną wartość zmiennej objaśnianej Y . Po prawej stronie równania (1.3) występują nieznanne wielkości $(\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_k)$, które należy oszacować. Noszą one nazwę **parametrów strukturalnych modelu**.

Ze składnikiem losowym modelu ekonometrycznego związane są parametry **struktury stochastycznej**. Są to parametry dotyczące rozkładu odchyłeń losowych modelu (np. wartość oczekiwana, wariancja odchyłeń losowych, współczynnik autokorelacji odchyłeń).

W notacji macierzowo-wektorowej liniowy model ekonometryczny zapisywany jest następująco:

$$\mathbf{y} = \mathbf{X}\mathbf{a} + \mathbf{e}, \quad (1.4)$$

gdzie: \mathbf{y} – wektor obserwacji zmiennej objaśnianej, \mathbf{X} – macierz obserwacji zmiennych objaśniających, \mathbf{a} – wektor ocen parametrów strukturalnych modelu, \mathbf{e} – wektor reszt;

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \quad \mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}.$$

Teoretyczne wartości zmiennej objaśnianej znajdujemy z równania macierzowego:

$$\hat{\mathbf{y}} = \mathbf{X}\mathbf{a}. \quad (1.5)$$

Podstawą badań ekonometrycznych powinna być teoria ekonomii. Jednakże nie zawsze w teorii ekonomii znajdujemy dostatecznie opracowane twierdzenia dające się weryfikować empirycznie. W związku z tym w analizach ekonometrycznych należy również wykorzystywać intuicję. Jak stwierdza B. Guzik „...ekonometryk musi być osobą, która umiejętnie i nieszablonowo łączy obie sfery działalności intelektualnej – naukę, czyli wiedzę, oraz sztukę, czyli fantazję. (...) Ani sama nauka, ani sama fantazja nie wystarczy” [Guzik, 2008, s. 18].

Podobną opinię prezentuje G.S. Maddala, stwierdzając, że „ekonometria to zastosowanie metod statystycznych i matematycznych do analizy danych ekonomicznych w celu nadania teoriom ekonomicznym kontekstu empirycznego oraz ich potwierdzenia lub odrzucenia” [Maddala, 2006, s. 31].

Warto nadmienić, że używana w publikacjach ekonometrycznych terminologia nie jest ujednolicona. Przykładowo, zmienne występujące w modelach ekonometrycznych noszą zazwyczaj nazwę zmiennych **objaśnianych** oraz **objaśniających**. Nazewnictwo to jest używane w celu podkreślenia, że celem modelu jest wyjaśnienie kształtowania się zmiennej Y . W matematyce mówi się natomiast o zmiennej **zależnej** i zmiennych **niezależnych**. W literaturze przedmiotu zmienną Y nazywa się niekiedy **regresorem**, a zmienną X – **regresatą** [Welfe, 2009, s. 27]. W zagadnieniach dotyczących prognozowania zmienną objaśnianą określa się mianem zmiennej prognozowanej (**predykaty**), a zmienne objaśniające – **predyktantami**. Zmienna wyjaśniana przez model nazywana jest zmienną **endogeniczną** (wewnętrzna), a zmienną niewyjaśnianą przez model – zmienną **egzogeniczną** (zewnętrzna). Terminy: zmienna zależna i zmienne niezależne są powszechnie stosowane w programach informatycznych.

Zasadniczym celem ekonometrii jest weryfikacja teorii ekonomicznych. Odbywa się ona w drodze konstrukcji dobrego modelu ekonometrycznego. Modelem ekonometrycznym nie jest dowolna funkcja, lecz taka, która jest dobrze uzasadniona teoretycznie (z punktu widzenia ekonomii i postulatów ekonometrycznych) oraz weryfikowalna empirycznie.

1.2. Etapy budowy modelu ekonometrycznego

Procedura budowy modelu ekonometrycznego ma wieloetapowy charakter. Model jest wynikiem postępowania łączącego w jedną całość metody matematyczne, statystyczne i informatyczne z wiedzą ekonomiczną i intuicją. W procesie budowy modelu ekonometrycznego można wyróżnić pięć etapów, a mianowicie [Strahl i in., 2004, s. 29–30]:

- specyfikację zmiennych,
- wybór analitycznej postaci modelu,
- estymację parametrów,
- weryfikację modelu,
- praktyczne wykorzystanie oszacowanego modelu.

Etap pierwszy obejmuje takie czynności, jak: **określenie celu** budowy modelu, ustalenie **listy zmiennych** (objaśnianej i objaśniających) oraz **zebranie informacji** statystycznych. Z reguły wiadomo od razu, co jest zmienną objaśnianą (wydajność pracy, popyt, produkcja itp.). W niektórych wypadkach zmienna objaśniana może mieć złożony charakter (np. poziom życia). Pojawia się wówczas problem, za pomocą jakich zmiennych scharakteryzować dane zjawisko złożone. W każdym razie wybór zmiennych powinien być dokonany tak, aby możliwie dokładnie została określona relacja między interesującymi nas procesami gospodarczymi. Ustalając listę potencjalnych zmiennych występujących w modelu koniecznym jest zwrócenie uwagi na siłę powiązań między nimi. Dobierając zmienne do modelu, kierujemy się wskazaniem teorii, wynikami innych badań, eksperymentami obliczeniowymi. Niejednokrotnie stosowane są tu statystyczne procedury doboru zmiennych objaśniających. Chodzi bowiem o to, by zmienne objaśniające gwarantowały najlepszy – z punktu widzenia przyjętego wskaźnika jakości – model. Przykładowo, często sugeruje się dobór takich zmiennych objaśniających, które mają wysoką zmienność bądź wykazują silne skorelowanie ze zmienną objaśnianą, a są słabo skorelowane ze sobą. Występujące w modelu ekonometrycznym zmienne powinny być jasno zdefiniowane i mieć wyraźną interpretację ekonomiczną.

Praktyczne wykorzystanie zbudowanego modelu ekonometrycznego jest – w znacznym stopniu – uzależnione od jakości zebranych danych statystycznych. Dane statystyczne, stanowiące podstawę konstrukcji modelu, mogą być specjalnie zbierane dla celów danego badania (**dane pierwotne**) lub pozyskiwane uprzednio do innych celów, ale wykorzystywane w procesie budowy modelu (**dane wtórne**). Dane statystyczne mogą mieć postać **szeregów czasowych, przekrojowych** lub **przekrojowo-czasowych**. Efektem specyfikacji zmiennych jest określona hipoteza modelowa, która będzie – w dalszych etapach konstrukcji modelu – podlegała weryfikacji.

Drugim etapem procedury modelowania ekonometrycznego jest **określenie postaci analitycznej modelu**. Teoria ekonomii nie dostarcza w tym względzie gotowych rozwiązań. Pomocne są tutaj wyniki innych badań, intuicja, eksperymenty obliczeniowe (próbujemy różnych funkcji, a spośród nich wybieramy tę, która najlepiej pasuje do danych empirycznych), wykresy przebiegu zmiennej objaśnianej względem zmiennych objaśniających, własności funkcji matematycznych (np. funkcja liniowa charakteryzuje się tym, że przyrostom zmiennej objaśniającej o jednostkę towarzyszy stały bezwzględny przyrost zmiennej objaśnianej, a w wypadku funkcji wykładniczej, stały jest przyrost względny), jak również testy statystyczne stosowane w odniesieniu do modeli liniowych *sensu stricte* lub modeli liniowych względem parametrów¹.

Trzecim etapem budowy modelu ekonometrycznego jest **estymacja** (szacowanie) parametrów strukturalnych. Podstawą tej czynności są zebrane informacje

¹ Por.: [Guzik, 2008, s. 25].

statystyczne. Informacje te powinny być wiarygodne i możliwie wyczerpujące. Liczba obserwacji powinna być przynajmniej 3–5-krotnie większa od liczby szacowanych parametrów modelu. Dobrze jest, jeśli obserwacje są **porównywalne**. Oznacza to, że wszystkie zmienne (objaśniana i objaśniające) powinny być wyrażone w tych samych jednostkach miary (w cenach stałych, jednostkach naturalnych, cenach bieżących itp.). Ważne znaczenie ma również ujednoczenie **wymiarów** zmiennych. Na przykład, jeśli Y jest zasobem, to zmienne X powinny mieć również charakter zasobów. Jeśli natomiast Y jest strumieniem, to zmienne objaśniające także powinny być strumieniami. **Zasoby** dotyczą wielkości danego zjawiska w ściśle określonym punkcie czasowym (np. według stanu na koniec miesiąca, roku), a **strumienie** powstają w wyniku sumowania wartości badanego zjawiska dla przedziałów o jednakowych rozpiętościach (np. roczna produkcja przedsiębiorstwa X jest sumą wartości produkcji z 12 miesięcy lub 4 kwartałów). O ile zatem zasoby mają wymiar W (stan na dany moment), o tyle strumienie – wymiar WT^{-1} (wielkość zjawiska w jednostce czasu). Rozróżnianie wymiarów zjawisk ma istotne znaczenie w analizie procesów społeczno-gospodarczych. Jak żartobliwie stwierdził M. Kalecki: „ekonomia jest nauką, w której ekonomisci stale mieszają pojęcia zasobów i strumieni, i popełniają wskutek tego błędy” [Lange, 1976, s. 547].

Za pomocą pojęć zasobu i strumienia definiowane są różne kategorie ekonomiczne. Na przykład proces produkcji można określić jako przekształcanie zasobów (pracy żywej i uprzedmiotowionej) w strumień wyrobów gotowych. Mamy tu bowiem zasoby W wydatkowane w czasie T , czyli $\frac{W}{T} = WT^{-1}$. Jeśli natomiast przekształcamy zysk (czyli strumień towarzyszący części wydatkowanej pracy żywej w procesie produkcji) w postać kapitału na dany moment (w zasób), to mamy do czynienia z działaniem typu $WT^{-1} \cdot T = W$, czyli z tworzeniem zasobów.

Do szacowania parametrów strukturalnych modelu najczęściej wykorzystywane są procedury aproksymacyjne. Polegają one na takim oszacowaniu, przy którym model najlepiej pasuje do danych empirycznych. Jest to spełnione wówczas, gdy minimalizowana jest rozbieżność między wartościami teoretycznymi (otrzymanymi z modelu) a odpowiednimi wartościami empirycznymi zmiennej Y . W praktyce, najczęściej stosowaną metodą estymacji nieznanymi parametrów strukturalnych $\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_k$ jest **klasyczna metoda najmniejszych kwadratów (KMNK)**.

Wektor ocen parametrów strukturalnych liniowego modelu ekonometrycznego otrzymany przy wykorzystaniu KMNK jest określony relacją:

$$\mathbf{a} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}. \quad (1.6)$$

Warunki stosowalności KMNK zostaną omówione w rozdziale trzecim.

Czwartym etapem konstrukcji modelu ekonometrycznego jest **weryfikacja**. Przebiega ona w dwóch płaszczyznach: ekonomicznej (merytorycznej) i statystycznej. Weryfikacja ekonomiczna oparta jest na badaniu zgodności modelu

z wiedzą o badanych zjawiskach i zdrowym rozsądkiem. W trakcie weryfikacji statystycznej wykorzystuje się m.in. odpowiednie mierniki dopasowania modelu ekonometrycznego do danych empirycznych, jak też sprawdza, czy np. zmienne objaśniające są istotne. Zwraca się też uwagę na sensowność poziomów i znaków ocen parametrów stojących przy zmiennych objaśniających. Na tym etapie bada się również **koicydentność** modelu. Mówimy, że model jest koicydentny, jeśli dla każdej zmiennej objaśniającej spełniony jest warunek:

$$\operatorname{sgn}(r_i) = \operatorname{sgn}(a_i) \quad \text{dla } i = 1, 2, \dots, k, \quad (1.7)$$

gdzie: a_i jest oszacowaniem parametru strukturalnego α_i występującego przy zmiennej objaśniającej X_i , natomiast r_i jest współczynnikiem korelacji liniowej Pearsona między zmienną X_i oraz Y . Symbol sgn oznacza znak (łac. *signum*). Model jest koicydentny, jeśli znaki przed wszystkimi ocenami parametrów są takie same, jak przy współczynnikach korelacji zmiennych objaśniających ze zmienną objaśnianą. Jeżeli model nie jest koicydentny, należy zmienić zbiór zmiennych objaśniających. Przyczynami braku koicydentności może być np. niewłaściwa postać analityczna modelu ekonometrycznego, bądź występująca współliniowość, tj. silna zależność między zmiennymi objaśniającymi.

Jeżeli wynik weryfikacji jest pozytywny, zbudowany model uznawany jest za dopuszczalny. W przeciwnym wypadku należy podjąć próbę skonstruowania innego modelu w drodze np. modyfikacji postaci analitycznej czy listy zmiennych objaśniających. Niekiedy może występować konieczność modyfikacji materiału statystycznego, przez jego rozszerzenie lub wyłączenie obserwacji nietypowych.

Problematyka weryfikacji modelu ekonometrycznego zostanie szerzej zaprezentowana w trzecim rozdziale tego podręcznika.

Rezultatem każdego badania ekonometrycznego powinno być skonstruowanie dobrego modelu. Tylko taki model może być bowiem wykorzystany praktycznie (**etap piąty**). Wykorzystanie modelu związane jest z celami jego budowy. Zgodnie z zamówieniem społecznym, inspirującym budowę modelu, może on być wykorzystany do prognozowania, symulacji, zwięzłego opisu zależności zmiennej objaśnianej od zmiennych objaśniających (cel analityczny) bądź też stanowić podstawę podejmowania decyzji kierowniczych.

Prognoza to kategoryczny sąd o wartości danego zjawiska w ustalonym momencie w przyszłości. Wartość logiczna tego sądu (prawda, fałsz) nie jest znana w momencie jego formułowania. Odmiennym od prognoz pojęciem są **symulacje**. Symulacje to sądy warunkowe, tzn. sądy typu: „jeśli zajdą dane okoliczności, to badana wielkość przyjmie następującą wartość” (determinacja czasowa nie musi tu wystąpić).

Operacyjnie rzecz biorąc, wykorzystanie modelu ekonometrycznego do prognozowania i symulacji polega na tym samym: należy obliczyć wartość zmiennej objaśnianej występującej w modelu przy ustalonych wartościach prognozowanych lub wariantowanych zmiennych objaśniających. Inna jest natomiast ich interpretacja. Na przykład, jeśli chcemy określić poziom sprzedaży przy następujących

wariantach indeksu dochodów: 1,4; 1,5; 1,6 – to obliczenia mają charakter symulacyjny. Jeśli natomiast interesuje nas wielkość sprzedaży artykułu X w roku następnym, w którym przewiduje się indeks dochodów równy 1,25 – to wyznaczamy prognozę.

1.3. Klasyfikacja modeli ekonometrycznych

Modele ekonometryczne mogą być klasyfikowane według różnych kryteriów:

- liczby zmiennych objaśnianych (lub liczby równań w modelu),
- roli czynnika czasu przy opisie modelowanych zjawisk,
- postaci analitycznej,
- charakteru poznawczego modelu,
- zakresu badania.

Kryterium podziału modeli ekonometrycznych ze względu na liczbę zmiennych objaśnianych jest równoznaczne z liczbą równań. Model opisujący jedną zmienną objaśnianą nazywamy **jednorównaniowym**, natomiast model opisujący kształtowanie się wielu zmiennych jednocześnie – **wielorównaniowym**. Przykładem modelu jednorównaniowego jest model opisujący popyt na określone dobro, zaś wielorównaniowego – modele rynków finansowych.

Ze względu na rolę czynnika czasu (własności dynamiczne) modele dzielimy na **dynamiczne** i **statyczne**. W modelach dynamicznych występują zmienne z opóźnieniami lub wyprzedzeniami czasowymi lub dowolna funkcja trendu. Wprowadzenie zmiennej z opóźnieniami oznacza, że bieżący – obserwowany w momencie t – stan zmiennej objaśnianej zależy od zaobserwowanych w przeszłości wartości zmiennych (tzn. w okresach $t - 1$, $t - 2$ itp.). Do klasy modeli dynamicznych zaliczane są również modele z parametrami zmieniającymi się w czasie. Jeśli model nie uwzględnia dynamiki analizowanych procesów – nazywamy go **statycznym** [Witkowska, 2005, s. 52].

Z punktu widzenia postaci analitycznej, modele dzielimy na **liniowe** i **nieliniowe**. W modelu liniowym zmienna objaśniana jest liniową funkcją zmiennych objaśniających i odchylenia losowego. Równanie modelu nieliniowego jest funkcją nieliniową. Wśród modeli nieliniowych możemy wyróżnić modele sprowadzalne do postaci liniowej i niesprowadzalne do postaci liniowej. W celu sprowadzenia modeli nieliniowych do liniowych wykorzystuje się odpowiednie transformacje (np. podstawianie, logarytmowanie). Mówimy wówczas o modelach **linearyzowalnych** [Dittmann, 2003, s. 132]. Istnieją również takie modele nieliniowe, które nie dają się sprowadzić do liniowych. Nazywamy je modelami nieliniowymi w ścisłym sensie. Przykładem modelu tego typu jest funkcja logistyczna.

Ze względu na wartości poznawcze modele ekonometryczne dzielimy na **przyczynowo-skutkowe**, **symptomatyczne**, **autoregresyjne** oraz modele **trendu** (tendencji rozwojowej) [Nowak, 1994, s. 8]. W modelach **przyczynowo-skutkowych** zjawiska (skutki) są wyjaśniane za pomocą innych, traktowanych jako

przyczyny. Przykładowo, wzrost zachorowań w okresie jesienno-zimowym powoduje zwiększenie popytu na leki przeciwdziałające przeziębieniu.

W modelach **symptomatycznych** przynajmniej niektóre spośród wyróżnionych zmiennych objaśniających nie stanowią bezpośredniej przyczyny dla zmiennej objaśnianej, ale są z nią silnie skorelowane. Na przykład poziom konsumpcji ryb w gospodarstwie domowym pozostaje w związku ze spożyciem mięsa i warzyw, chociaż spożycia wymienionych produktów nie pozostają ze sobą w związku przyczynowym.

Modele **autoregresyjne** to takie modele, w których jedyną zmienną objaśniającą jest zmienna objaśniana o **opóźnionych w czasie wartościach**. Tego rodzaju modele mają zastosowanie głównie do zjawisk odznaczających się inercją.

Modele **trendu** (tendencji rozwojowej) opisują rozwój zjawisk w czasie. W modelach tych jedyną zmienną objaśniającą jest czas. Zmienna czasowa (t) przybiera wartości kolejnych liczb naturalnych przyporządkowanych jednostkom czasu badanego okresu.

Kryterium zakresu badania pozwala sklasyfikować modele na **mikro-, mezo- i makroekonomiczne**. Modele mikroekonomiczne są budowane dla procesów zachodzących w przedsiębiorstwach. Przykładem modeli w skali mezo- są m.in. modele regionalne oraz modele sektorów gospodarki. W analizie zjawisk obserwowanych w skali całej gospodarki, rynków międzynarodowych lub gospodarki światowej – budowane są modele makroekonomiczne.

Rozdział 2. Dobór zmiennych objaśniających do modelu ekonometrycznego

2.1. Współczynnik zmienności jako kryterium doboru zmiennych

Kryterium wyboru zmiennych objaśniających do modelu ekonometrycznego powinna być merytoryczna znajomość badanego zjawiska. Chodzi tu bowiem o taki wybór zmiennych, które mają istotny wpływ na kształtowanie się zmiennej objaśnianej. Takie zebrane zmienne tworzą zbiór **potencjalnych** zmiennych objaśniających. Wysoka liczebność tego zbioru może stwarzać określone trudności w zebraniu materiału statystycznego, stanowiącego podstawę konstrukcji modelu ekonometrycznego. W celu określenia **właściwego** zestawu zmiennych objaśniających stosuje się zazwyczaj trzy podejścia oparte na [Witkowska, 2005, s. 30]:

- wiedzy wynikającej z teorii ekonomii,
- zebranych materiale statystycznym, doświadczeniu i intuicji (jeśli teoria ekonomii nie jest dostatecznie rozwinięta),
- metodach statystycznych (gdy teoria ekonomiczna i analiza empiryczna połączone z doświadczeniem nie są wystarczające).

M. Gruszczyński podaje, że dobór zmiennych objaśniających do modeli powinien uwzględniać przedstawione poniżej zasady [Gruszczyński, 2010, s. 79–80].

1. Zmienne powinny być merytorycznie (teoretycznie) uzasadnione, tzn. wyjaśniać kształtowanie się zmiennej objaśnianej Y . W większości wypadków nie mamy takich teorii, a tylko określone intuicje, które uznawane są za pewne wskazania merytoryczne.
2. Liczba zmiennych objaśniających powinna być – nawet w wypadku posiadania dużych zbiorów danych – ograniczona. Jest to zasada selekcji i nie mnożenia liczby zmiennych objaśniających.
3. Ograniczoną liczbę zmiennych objaśniających należy wybierać z różnych ich grup. Grupami tymi mogą być np. warunki ekonomiczne rodziny, stan demograficzny, poglądy polityczne. Każda z tych grup zawiera (reprezentuje) wiele zmiennych. Zmienne opisujące kształtowanie się zmiennej objaśnianej powinny być wybrane ze wszystkich grup. Ograniczając liczbę zmiennych objaśniających, należy kierować się ich reprezentatywnością.
4. Zmienne, które znalazły się w zbiorze wybranych zmiennych objaśniających nie powinny charakteryzować się endogenicznością. Endogeniczność ozna-

cza, iż dana zmienna objaśniająca modelu jest skorelowana ze składnikiem losowym. Podkreślić należy, że wybór zmiennych pozbawionych własności endogeniczności opiera się raczej na intuicji niż na teorii.

5. W modelu ekonometrycznym powinny znaleźć się zmienne istotne statystycznie, tzn. w uzasadniony statystycznie sposób mające związek ze zmienną objaśnianą.
6. Zmienne wybrane do modelu nie powinny wykazywać współliniowości. Współliniowość oznacza wysoką korelację między zmiennymi objaśniającymi. Utrudnia to, a niekiedy wręcz uniemożliwia wydzielenie indywidualnego wpływu każdej ze zmiennych objaśniających na zmienną objaśnianą.

W pierwszym kroku z potencjalnego zbioru zmiennych objaśniających należy wyeliminować te, których zróżnicowanie nie jest wystarczające. Noszą one nazwę zmiennych **quasi-stałych**. Miernikiem zmienności (zróżnicowania) zmiennej jest współczynnik zmienności określony wzorem:

$$V_i = \frac{s_i}{\bar{x}_i}, \quad (2.1)$$

gdzie: s_i – odchylenie standardowe zmiennej X_i , \bar{x}_i – średnia arytmetyczna zmiennej X_i .

Po ustaleniu przez badacza wartości krytycznej V^* , przyrównuje się ją do wartości współczynnika zmienności dla danej zmiennej. Zmienne spełniające nierówność:

$$V_i \leq V^* \quad (2.2)$$

uznaje się za quasi-stałe i eliminuje ze zbioru potencjalnych zmiennych objaśniających. Zmienne quasi-stałe posiadają zbyt niski poziom zmienności i nie wnoszą istotnych informacji o kształtowaniu się zmiennej objaśnianej. Wartość krytyczna V^* jest ustalana arbitralnie, np. na poziomie 0,10 lub 0,15. Zależy to od warunku przyjętego przez osobę tworzącą model ekonometryczny.

Przykład 2.1. Mamy dokonać wyboru zmiennych objaśniających do liniowego modelu ekonometrycznego opisującego kształtowanie się produkcji żywca w decytonach na 1 ha użytków rolnych (Y) w dużych gospodarstwach rolnych pewnego regionu. Potencjalny zbiór zmiennych objaśniających składa się z sześciu elementów, a mianowicie:

- x_1 – plony owsa w dt/ha,
- x_2 – plony jęczmienia w dt/ha,
- x_3 – plony żyta w dt/ha,
- x_4 – zbiory ziemniaków w dt/ha,
- x_5 – zbiory siana łąkowego w t/ha,
- x_6 – średnie ceny skupu żywca w zł/kg.

Liczbowe wartości wyróżnionych zmiennych, obejmujące ostatnie dziesięć lat, przedstawiono w tabeli 2.1.

Do wyboru zmiennych należy wykorzystać metodę opartą o współczynniki zmienności, przy przyjęciu krytycznej wartości na poziomie $V^* = 0,15$.

Tabela 2.1. Zaobserwowane wartości zmiennych

Lata	y_t	Potencjalne zmienne objaśniające					
		x_1	x_2	x_3	x_4	x_5	x_6
1	5	10	12	8	30	6	2,6
2	5	12	14	8	31	7	2,7
3	6	12	14	9	32	7	2,9
4	6	15	15	9	34	7	3,0
5	8	15	18	10	34	8	3,5
6	8	17	20	10	36	8	3,5
7	9	18	20	10	37	8	3,7
8	10	19	21	11	38	8	3,8
9	10	20	22	11	38	9	4,0
10	11	24	23	11	40	9	4,3

Źródło: [Grabowski, 2002, s. 24].

Średnie arytmetyczne potencjalnych zmiennych objaśniających są równe:

$$\bar{x}_1 = \frac{162}{10} = 16,2; \quad \bar{x}_2 = \frac{179}{10} = 17,9; \quad \bar{x}_3 = \frac{97}{10} = 9,7;$$

$$\bar{x}_4 = \frac{350}{10} = 35; \quad \bar{x}_5 = \frac{77}{10} = 7,7; \quad \bar{x}_6 = \frac{34}{10} = 3,4.$$

Odchylenia standardowe dla poszczególnych zmiennych objaśniających wynoszą:

$$s_1 = \sqrt{\frac{163,6}{10}} = 4,04; \quad s_2 = \sqrt{\frac{134,90}{10}} = 3,67; \quad s_3 = \sqrt{\frac{12,1}{10}} = 1,1;$$

$$s_4 = \sqrt{\frac{100}{10}} = 3,16; \quad s_5 = \sqrt{\frac{6,41}{10}} = 0,80; \quad s_6 = \sqrt{\frac{2,98}{10}} = 0,55.$$

Obliczone według wzoru (2.1) współczynniki zmienności poszczególnych zmiennych V przyjmują wartości:

$$V_1 = \frac{4,04}{16,2} = 0,25; \quad V_2 = \frac{3,67}{17,9} = 0,21; \quad V_3 = \frac{1,1}{9,7} = 0,11;$$

$$V_4 = \frac{3,16}{35} = 0,09; \quad V_5 = \frac{0,8}{7,7} = 0,10; \quad V_6 = \frac{0,55}{3,4} = 0,16.$$

Okazuje się, że współczynniki dla zmiennych objaśniających: x_3, x_4 oraz x_5 są niższe od założonej wartości krytycznej 0,15. Zmienne te są więc quasi-stałymi, które należy wyeliminować ze zbioru potencjalnych zmiennych objaśniających. W zbiorze zmiennych objaśniających modelu ekonometrycznego powinny znaleźć się zmienne: plony owsa (x_1), plony jęczmienia (x_2) oraz średnie ceny skupu żywca (x_6).