MODELE EKONOMETRYCZNE

Model ekonometryczny to opis stochastycznej zależności badanego zjawiska ekonomicznego od czynników kształtujących go, wyrażony w postaci równości lub układu równości.

Jeśli np. rozpatrujemy zjawisko popytu na określony towar lub grupę towarów i przyjmiemy, że głównym czynnikiem kształtującym popyt jest cena to możemy rozpatrywać model

D = f(P) D - popyt, P - cena.

Z prawa malejącego popytu wynika, że funkcja f powinna być malejąca ($(P_1 < P_2 \Rightarrow f(P_1) > f(P_2))$. Zależność tę możemy zrealizować za pomocą różnych funkcji malejących, najprostsza z nich to funkcja liniowa:

$$D = a + bP$$

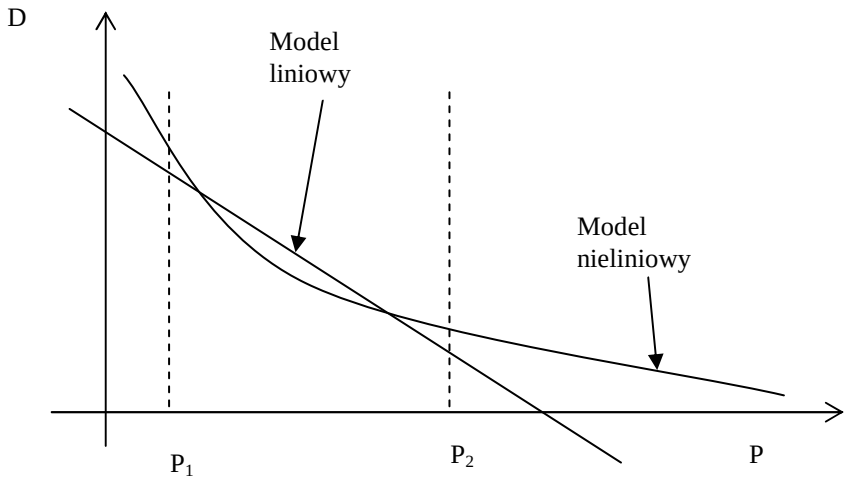
(liniowy model popytu), a > 0; b < 0

jeśli model liniowy nie pasuje do zaobserwowanych wielkości to należy zastosować model nieliniowy np. model potęgowy:

$$D = a \cdot P^b$$

(potęgowy model popytu), $a > 0$; $b < 0$

Dla pewnych zakresów cen model liniowy może być dobrym przybliżeniem modelu nieliniowego



Niekiedy model z jedną zmienną źle opisuje badane zjawisko, wtedy możemy rozpatrywać model z wieloma zmiennymi.

W modelu popytu drugim czynnikiem kształtującym popyt może być dochód, wtedy rozpatrujemy zależność:

D = f(P, I) I - dochód ludności.

Zależność tę możemy jak poprzednio zrealizować za pomocą funkcji liniowej

$$D = a + bP + cI$$

lub potęgowej

$$D = a \cdot P^b I^c$$

Ogólna postać modelu w postaci jednej równości:

$$Y = f(X, \varepsilon)$$

X, Y - zmienne ,

(X może być postaci X = (X_1, X_2, ..., X_k)),

ε -element losowy

Powody uwzględniania elementu losowego w modelu ekonometrycznym:

– nie uwzględnienie wszystkich czynników kształtujących badane zjawisko (najczęściej nie uwzględniamy czynników mających mały wpływ i element losowy reprezentuje łączny wpływ takich zmiennych),

– możliwość występowania błędów w pomiarze wielkości zmiennych,
– brak pewności czy przyjęta do obliczeń postać funkcyjna modelu jest prawidłowa.

Uproszczona **klasyfikacja zmiennych** w modelu

– zmienna endogeniczna – zmienna, której wartości określone są w modelu,
– zmienna egzogeniczna – zmienna, której wartości określone są poza modelem,
– zmienna objaśniana – występuje po lewej stronie równań modelu,
– zmienna objaśniająca – występuje po prawej stronie równań modelu.

Każda ze zmiennych może być bieżąca lub opóźniona.

Uwaga:
W modelach wielowymiarowych zmienna objaśniana może być jednocześnie zmienną objaśniającą.

Przykład 1.

Rozpatrzmy model wzrostu gospodarczego

$$\begin{cases} DN_t = aNI_{t-4}^{b} Z_t^{c} \varepsilon_{1t} \\ NI_t = dDN_t + \varepsilon_{2t} \end{cases}$$

gdzie

DN - dochód narodowy,

NI - nakłady inwestycyjne,

Z - zatrudnienie,

a,b,c,d - parametry strukturalne,

$\varepsilon_1, \varepsilon_2$ - elementy losowe

Klasyfikacja:

- zmienne endogeniczne: DN_t, NI_t, NI_{t-4}
- zmienne egzogeniczne: Z_t
- zmienne objaśniane: DN_t, NI_t
- zmienne objaśniające: NI_{t-4}, Z_t, DN_t
- zmienne bieżące: DN_t, Z_t, NI_t
- zmienne opóźnione: NI_{t-4} .

Klasyfikacja modeli

Modele klasyfikujemy ze względu na następujące kryteria:

a) liczba zależności w modelu

- modele jednorównaniowe,
- modele wielorównaniowe,

b) postać zależności funkcyjnej,

- modele liniowe,
- modele nieliniowe (potęgowe, wykładnicze, itp.).

c) rola czasu w równaniach,

- modele statyczne (nie uwzględniają czasu),
- modele dynamiczne.

Przykład 2

Model z przykładu 1 jest:

- dwurównaniowy,
- nieliniowy,
- dynamiczny.

Jednorównaniowy model liniowy z jedną zmienną objaśniającą

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

gdzie:

Y - zmienna objaśniana, y_i - wartości (obserwacje) zmiennej Y; $i = 1,...,n$ - numer obserwacji,

X - zmienna objaśniająca, x_i - wartości zmiennej X,

β_0, β_1 - parametry strukturalne (ich przybliżoną wartość wyznacza się na podstawie obserwacji (x_i, y_i))

ε - składnik losowy.

Zakładamy, że

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \qquad i = 1,2,...,n$$

tzn. każda zaobserwowana wartość y_i jest funkcją liniową x_i z dokładnością do składnika losowego ε_i .

Zakładamy również, że x_i są ustalonymi wartościami (nielosowymi), takimi samymi w powtarzalnych próbach. Składniki losowe ε_i są losowymi zmiennymi niezależnymi o zerowej

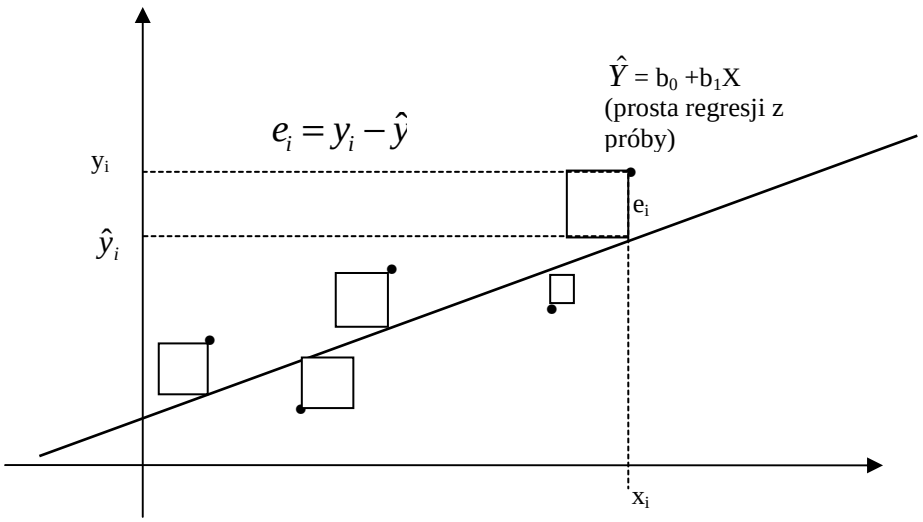
wartości przeciętnej i wariancji, która nie zależy od i (homoskedastyczność).

Aby wyznaczyć przybliżoną wartość parametrów strukturalnych β_0, β_1 na podstawie próby stosujemy metodę najmniejszych kwadratów (MNK).

MNK polega na wyznaczeniu takich przybliżeń

$$b_0 \approx \beta_0 \qquad b_1 \approx \beta_1$$

aby dla danych obserwacji (x_i, y_i) suma kwadratów odchyleń zaobserwowanych wartości y_i od wartości teoretycznych $\hat{y}_i = \beta_0 + \beta_1 x_i$ była minimalna, tzn. chcemy wyznaczyć minimum funkcji:



$$(*) \qquad S(b_0, b_1) = \sum_{i=1}^{n} e_i^{\,2} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i)^2$$

$e_i = y_i - \hat{y}_i$ nazywamy **resztami** modelu regresji

Należy wyznaczyć prostą regresji tak aby suma pól kwadratów była minimalna.
Obliczając pochodne cząstkowe funkcji (*) i przyrównując do zera otrzymujemy
(układ równań normalnych)

$$\frac{\partial S}{\partial \beta_0} = 2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)(-1) = -2\left(\sum_{i=1}^{n} y_i - \beta_1 \sum_{i=1}^{n} x_i - n\beta_0\right) = 0$$

$$\frac{\partial S}{\partial \beta_1} = 2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)(-x_i) = -2\left(\sum_{i=1}^{n} y_i x_i - \beta_1 \sum_{i=1}^{n} x_i^2 - \beta_0 \sum_{i=1}^{n} x_i\right) = 0$$

rozwiązując otrzymany układ równań otrzymamy wzory na przybliżone wartości parametrów strukturalnych

$$b_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2} = \frac{\frac{1}{n}\sum x_i y_i - \overline{xy}}{\frac{1}{n}\sum x_i^2 - (\overline{x})^2} = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sum (x_i - \overline{x})^2}$$

$$b_0 = \overline{y} - b_1 \overline{x}$$

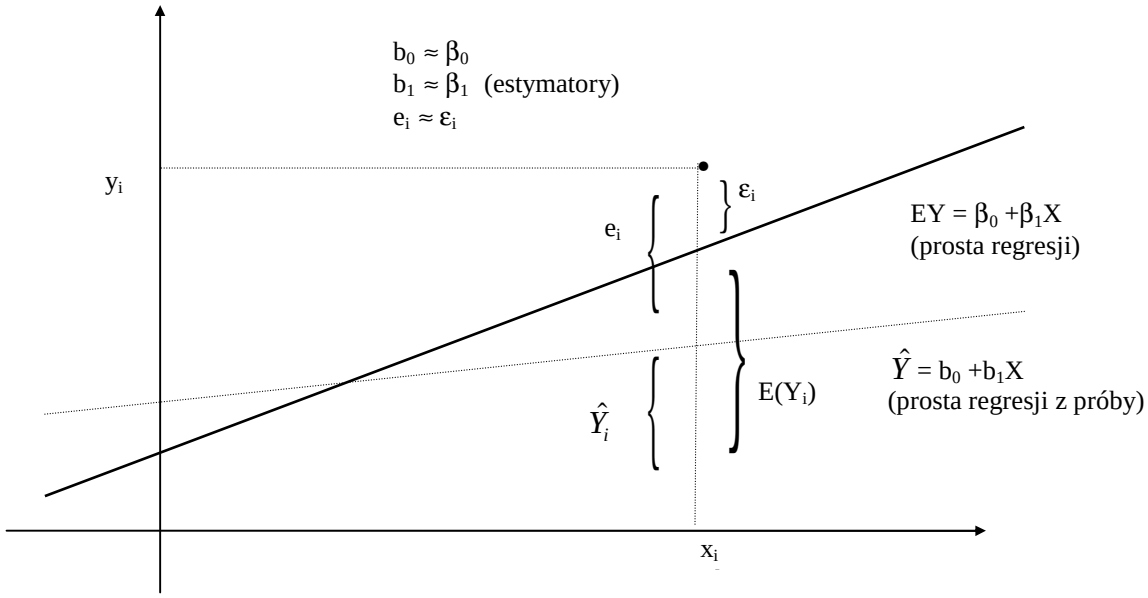
Prostą

$$\hat{Y} = b_0 + b_1 X$$

nazywamy **prostą regresji z próby**.

Model regresji liniowej:

Uwaga



Miary dopasowania.

Wariancja resztowa:

Wariancja resztowa to uśrednienie pól kwadratów zbudowanych na resztach
i odzwierciedla stopień dopasowania prostej regresji do danych statystycznych.

Niech, $e_i = y_i - \hat{y}_i$, gdzie $\hat{y}_i = b_0 + b_1 x_i$ wtedy

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

czyli

$$S_e^2 = \frac{\sum y_i^2 - b_0 \sum y_i - b_1 \sum x_i y_i}{n-2}$$

$S_e = \sqrt{S_e^2}$ oznacza średnie (standardowe) odchylenie od prostej regresji.

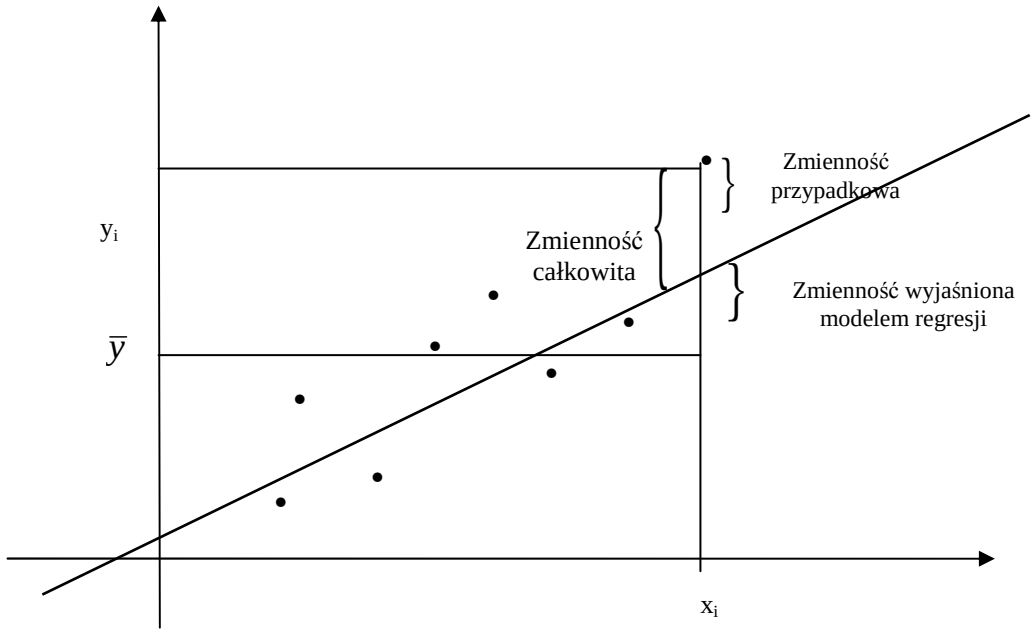
Dopasowanie modelu do danych empirycznych można oceniać odchyleniem standardowym reszt lecz jest to miara bezwzględna i nieunormowana, dlatego do porównań lepsze są miary względne lub unormowane.

Najprostszą względną miarą dopasowania jest **współczynnik zmienności losowej** :

$$V_e = \frac{S_e}{\overline{Y}} 100\%$$

Współczynnik ten informuje jaką część średniej wartości badanego zjawiska stanowi odchylenie standardowe reszt.

Mniejsze wartości tego współczynnika wskazują na lepsze dopasowanie modelu do danych empirycznych, niekiedy żąda się aby np. $V_e < 0{,}2$.



Wprowadzamy oznaczenia:

Całkowita suma kwadratów (zmienność całkowita): $CSK = \sum (y_i - \bar{y})^2$

Wyjaśniona suma kwadratów (zmienność wyjaśniona): $WSK = \sum (\hat{y}_i - \bar{y})^2$

Niewyjaśniona suma kwadratów (zmienność przypadkowa): $NSK = \sum e_i^2$

$gdzie: \hat{y}_t = b_0 + b_1 x_t$

Własność: $\sum (y_i - \bar{y})^2 = \sum (\hat{y}_i - \bar{y})^2 + \sum e_i^2$

Czyli CSK = WSK + NSK

Miarą dopasowania modelu do rzeczywistości (wartości zaobserwowanych) jest również współczynnik determinacji R^2

Współczynnik determinacji:

$$R^2 = \frac{WSK}{CSK} \qquad R^2 \in \langle 0, 1 \rangle$$

współczynnik ten określa jaka część całkowitej zmienności zmiennej objaśnianej została wyjaśniona przez model regresji liniowej.

$$R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2} = 1 - \frac{\sum e_i^2}{\sum(y_i - \bar{y})^2} = \frac{b_0 \sum y_i + b_1 \sum x_i y_i - n\bar{y}^2}{\sum y_i^2 - n(\bar{y})^2} =$$

$$= \frac{b_1\left(\sum x_i y_i - n\bar{x}\bar{y}\right)}{\sum y_i^2 - n(\bar{y})^2} = \frac{\mathrm{cov}^2(X,Y)}{S_X^2 S_Y^2} = r^2$$

Przykład

Badano zależności kosztów całkowitych (w tys. zł.) Y od wielkości produkcji (tys. szt.) X w 6-ciu zakładach produkcyjnych.

x_t	2	4	3	2	6	1
y_t	2	5	4	4	7	2

Dla modelu $Y = \beta_0 + \beta_1 x + \varepsilon$ wyznaczamy przybliżone wartości parametrów strukturalnych i współczynnik determinacji.

x_i	y_i	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$\hat{y}_i$	$\hat{y}_i - \bar{y}$	$(\hat{y}_i - \bar{y})^2$
2	2	-1	-2	2	1	4	3	-1	1
4	5	1	1	1	1	1	5	1	1
3	4	0	0	0	0	0	4	0	0
2	4	-1	0	0	1	0	3	-1	1
6	7	3	3	9	9	9	7	3	9
1	2	2	-2	4	4	4	2	-2	4
18	24	0	0	16	16	18	24	0	16

$$\bar{x} = \frac{18}{6} = 3; \qquad \bar{y} = \frac{24}{6} = 4; \qquad b_1 = \frac{16}{16} = 1; \quad b_0 = 4 - 1 * 3 = 1$$

zatem związek pomiędzy kosztami całkowitymi a wielkością produkcji wyraża się zależnością liniową w postaci

$$\hat{Y} = 1 + X$$

Współczynnik determinacji

$$R^2 = \frac{16}{18} = 0{,}89$$

należy oczekiwać, że rozpatrywany model wyjaśnia 89% całkowitej zmienności badanego zjawiska.

JEDNORÓWNANIOWY MODEL LINIOWY - POSTAĆ OGÓLNA.

Ogólna postać modelu liniowego z k zmiennymi objaśniającymi.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon$$

gdzie:

Y - zmienna objaśniana

X_i - zmienne objaśniające, i = 1, 2, ..., k

β_i - parametry strukturalne, i = 0, 1, 2, ..., k

ε - składnik losowy.

Założenia:

Niech

n - liczba obserwacji,

t = 1, 2, ..., n - numery obserwacji,

y_t , ε_t , x_{t1} , x_{t2} , ..., x_{tk} , - zaobserwowane wartości

zmiennej objaśniającej, składnika losowego

i zmiennych objaśniających , t = 1, 2, ..., n.

Założenie 1.

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + ... + \beta_k x_{tk} + \varepsilon_t$$

dla t = 1, 2, ..., n

Założenie 2.

a) x_{t1} , x_{t2} , ..., x_{tk} , - wartości ustalone (nie są losowe),

b) x_{t1} , x_{t2} , ..., x_{tk} , - liniowo niezależne,

c) n > k + 1

Założenie 3.

ε_t - są **niezależnymi** zmiennymi losowymi o jednakowym rozkładzie prawdopodobieństwa, N(0, σ).

$$E(\varepsilon_t) = 0,$$
$$D^2(\varepsilon_t) = \sigma^2$$

(homoscedastyczność)

Zapis modelu w postaci macierzowej.
Niech:

$$Y=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad X=\begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \quad \beta=\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad \varepsilon=\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Zatem powyższy model można zapisać w postaci.

$$Y = X\beta + \varepsilon$$

Stosując metodę najmniejszych kwadratów otrzymujemy wektor estymatorów parametrów strukturalnych $b \approx \beta$:

Metoda najmniejszych kwadratów (MNK).

Rozpatrujemy funkcję:

$$S(b)=\sum (y_i - b_0 - x_{i1}b_1 - x_{i2}b_2 - \ldots - x_{ik}b_k)^2 =$$

$$=(y-Xb)^T(y-Xb)=y^Ty-2b^TX^Ty+b^TX^TXb$$

Uwaga .

Reguły różniczkowania względem wektora są następujące:

$$\frac{\partial b^T a}{\partial b}=a \qquad \frac{\partial (b^TAb)}{\partial b}=Ab+A^Tb$$

i przyrównujemy do zera.

$$-2X^Ty+2X^TXb=0$$

stąd

$$\boxed{b=\left(X^TX\right)^{-1}X^TY}$$

Nieobciążoność estymatora b.

$$b = (X^T X)^{-1} X^T (X\beta + \varepsilon) =$$

$$= (X^T X)^{-1} X^T X\beta + (X^T X)^{-1} X^T \varepsilon =$$

$$= \beta + (X^T X)^{-1} X^T \varepsilon$$

zatem

$$Eb = \beta + (X^T X)^{-1} X^T (E\varepsilon) = \beta$$

Macierz kowariancji dla b:

$$\operatorname{cov} b = E[(b - Eb)(b - Eb)^T] =$$

$$= E[(X^T X)^{-1} X^T \varepsilon\varepsilon^T X (X^T X)^{-1}] =$$

$$= (X^T X)^{-1} X^T E(\varepsilon\varepsilon^T) X (X^T X)^{-1} =$$

$$= (X^T X)^{-1} X^T X (X^T X)^{-1} \sigma^2 = (X^T X)^{-1} \sigma^2$$

Twierdzenie (Gaussa – Markowa).
Najlepszym, nieobciążonym, liniowym estymatorem wektora β w modelu liniowym, w którym $E\varepsilon = 0$ oraz $E\varepsilon\varepsilon^T = I\sigma^2$, jest estymator b uzyskany metodą najmniejszych kwadratów.

Wektor reszt uzyskanych z równania regresji oszacowanego metodą najmniejszych kwadratów jest równy:

$$e = y - Xb = X\beta + \varepsilon - X[\beta + (X^T X)^{-1} X^T \varepsilon] =$$

$$= [I - X(X^T X)^{-1} X^T]\varepsilon$$

Nieobciążonym estymatorem parametru σ^2 jest **wariancja resztowa**:

Niech, $e = Y - \hat{Y}$, gdzie $\hat{Y} = Xb$ wtedy

$$S_e^2 = S^2 = \frac{e^T e}{n-(k+1)} = \frac{\sum_{i=1}^{n} e_i^2}{n-(k+1)} = \frac{1}{n-(k+1)}\left(Y^T Y - b^T X^T Y\right)$$

dla k = 1

$$S_e^2 = \frac{\sum y_i^2 - b_0 \sum y_i - b_1 \sum x_i y_i}{n-2}$$

Miary dopasowania.

Współczynnik determinacji:

$$R^2 = \frac{b^T X^T Y - n\bar{Y}^2}{Y^T Y - n\bar{Y}^2}$$

dla k = 1

$$R^2 = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2} = 1 - \frac{\sum e_i^2}{\sum (y_i - \bar{y})^2} =$$

$$= \frac{b_0 \sum y_i + b_1 \sum x_i y_i - n\bar{y}^2}{\sum y_i^2 - n(\bar{y})^2} =$$

$$= \frac{b_1\left(\sum x_i y_i - n\overline{xy}\right)}{\sum y_i^2 - n(\bar{y})^2} = \frac{\operatorname{cov}^2(X,Y)}{S_X^2 S_Y^2}$$

Współczynnik zbieżności:

$$\Phi^2 = 1 - R^2 = \frac{Y^T Y - b^T X^T Y}{Y^T Y - n\overline{Y}^2}$$

Skorygowany współczynnik zbieżności:

$$\hat{\Phi}^2 = \Phi^2 \frac{n-1}{n-(k+1)}$$

Skorygowany współczynnik determinacji:

$$\hat{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-(k+1)}$$

Standardowe błędy oszacowania parametrów strukturalnych.

Rozpatrujemy macierz $S_e^2\left(X^T X\right)^{-1}$

na głównej przekątnej tej macierzy mamy wariancje tych błędów tzn. $S^2(b_i)$, i = 0, 1, ..., k.

Zatem

$$\boxed{S(b_i) = \sqrt{S^2(b_i)}}, \qquad \text{i = 0, 1, ..., k.}$$

dla k = 1

$$\boxed{S(b_1) = \frac{S_e}{\sqrt{\sum (x_i - \bar{x})^2}}}$$

$$\boxed{S(b_0) = \frac{S_e\sqrt{\sum x_i^2}}{\sqrt{n\sum (x_i - \bar{x})^2}} = S(b_1)\cdot\sqrt{\frac{1}{n}\sum x_i^2}}$$

Stosujemy niekiedy zapis $\hat{Y} = \underset{(S(b_0))}{b_0} + \underset{(S(b_1))}{b_1} X$

Przykład.

Y - wydajność pracy (mln zł/zatr.),

X_1 - techniczne uzbrojenie pracy (mln zł/zatr.),

X_2 - zatrudnienie (setki osób),

Rozpatrujemy model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$.

i	1	2	3	4	5	6	7	8	9
Y	10	9	11	13	12	15	14	16	17
X_1	0,6	0,5	0,9	1,1	1,0	1,2	0,9	1,3	1,5
X_2	10	8	8	9	8	7	5	4	4

$$X^T Y = \begin{bmatrix} 117 \\ 123{,}6 \\ 780 \end{bmatrix} \qquad Y^T Y = 1581$$

$$\left(X^T X\right)^{-1} = \begin{bmatrix} 7{,}9687666 & -3{,}8636363 & -0{,}5705741 \\ -3{,}8636363 & 2{,}2727272 & 0{,}2272727 \\ -0{,}5705741 & 0{,}2272727 & 0{,}0490430 \end{bmatrix}$$

wyznaczymy $b,\ R^2,\ S(b_i)\ \ i = 0, 1, 2$.

Rozwiązanie:

$$b = \left(X^T X\right)^{-1} X^T Y = \begin{bmatrix} 7{,}9687666 & -3{,}8636363 & -0{,}5705741 \\ -3{,}8636363 & 2{,}2727272 & 0{,}2272727 \\ -0{,}5705741 & 0{,}2272727 & 0{,}0490430 \end{bmatrix} \begin{bmatrix} 117 \\ 123{,}6 \\ 780 \end{bmatrix} = \begin{bmatrix} 9{,}752 \\ 6{,}136 \\ -0{,}413 \end{bmatrix};$$

zatem równanie płaszczyzny regresji ma postać:

$$\hat{Y} = 9{,}752 + 6{,}136 X_1 - 0{,}413 X_2$$

Wariancja resztowa jest równa $S_e^2 = 0{,}572$

stąd $S_e = 0{,}756$;

Błędy standardowe estymatorów parametrów strukturalnych:

$S(b_0) = 2{,}135$;

$S(b_1) = 1{,}140$

$S(b_2) = 0{,}168$

Uwzględniamy je w zapisie:

$$\hat{Y} = 9{,}752 + 6{,}136 X_1 - 0{,}413 X_2$$
$$(2{,}135) \quad (1{,}140) \quad (0{,}168)$$

Współczynnik determinacji wynosi

$$R^2 = 0{,}943$$

Przedziały ufności dla β_i , i = 0, 1, 2, ..., k ;

dla poziomu ufności 1-α mamy:

$$\boxed{\beta_i \in \langle b_i - u_\alpha S(b_i); b_i + u_\alpha S(b_i) \rangle}$$

gdzie u_α odczytujemy z tablicy rozkładu Studenta: $P\left(\left|T_{n-(k+1)}\right| > u_\alpha\right) = \alpha$.

Weryfikacja hipotez dla β_i , i = 0, 1, 2, ..., k ;
dla poziomu istotności α rozpatrzymy dwa testy:

1) Uogólniony test Walda.

Wysuwamy dwie hipotezy:

$$H_0\left(\beta_1 = \beta_2 = ... = \beta_k = 0\right)$$

$$H_1\left(\text{co najmniej jedno } \beta_i \neq 0,\ i = 1, 2, \ ...k\right)$$

Stosujemy statystykę

$$U_n = \frac{R^2}{1-R^2}\frac{n-(k+1)}{k} = \frac{b^T X^T Y - \frac{\left(\mathbf{1}^T Y\right)^2}{n}}{e^T e}\frac{n-(k+1)}{k}$$

Rozpatrujemy zbiór krytyczny:

$$K = < k;+\infty)$$

gdzie k odczytujemy dla poziomu istotności
α z tablicy rozkładu Fiszera-Snedecora
dla (k, n - (k + 1)) stopni swobody.
Decyzje:
Jeśli $U_n \in K$ to H_0 odrzucamy ,
Jeśli $U_n \notin K$ to nie ma podstaw do odrzucenia H_0 .

2) Test dla poszczególnych parametrów β_i , i = 0, 1, 2, ..., k.

Wysuwamy dwie hipotezy:

$H_0\left(\beta_i = \beta_i^0\right)$, H_1 - jedną z trzech poniższych hipotez.

Rozpatrujemy statystykę i zbiór krytyczny wg tabeli:

H_1	Statystyka	Zbiór krytyczny	Odczyt k
$\beta_i \neq \beta_i^0$	$U_n = \frac{b_i - \beta_i^0}{S(b_i)}$	$K = (-\infty;-k > \cup < k;+\infty)$	$P\left(\left\|T_{n-(k+1)}\right\| > k\right) = \alpha$
$\beta_i > \beta_i^0$		$K =< k;+\infty)$	$P\left(\left\|T_{n-(k+1)}\right\| > k\right) = 2\alpha$
$\beta_i < \beta_i^0$		$K = (-\infty;-k >$	$P\left(\left\|T_{n-(k+1)}\right\| > k\right) = 2\alpha$

Decyzje:

Jeśli $U_n \in K$ to H_0 odrzucamy ,

Jeśli $U_n \notin K$ to nie ma podstaw do odrzucenia H_0 .

Uwaga

Jeśli badamy istotność parametru β_i to przyjmujemy $\beta_i^0 = 0$ tzn. rozpatrujemy hipotezę

$$H_0\left(\beta_i = 0\right)$$

Badanie losowości reszt (test serii).

Rozpatrujemy hipotezy
H_0(reszty modelu mają charakter losowy),
H_1(reszty modelu nie mają charakteru losowego),

Resztom przypisujemy symbol *a* lub *b*:
a - gdy $e_i > 0$,
b - gdy $e_i < 0$
(reszt $e_i = 0$ nie rozpatrujemy).

Serie to podciągi złożone z jednakowych symboli.

Stosujemy statystykę:

$$U_n = \text{liczba serii}$$

Zbiór krytyczny:

$$K = (0; k>$$

gdzie k odczytujemy z tablicy dla poziomu istotności α i liczb n_1 oraz n_2 , gdzie
n_1 - liczba symboli *a*,
n_2 - liczba symboli *b*,
Tablica dla $\alpha = 0{,}05$:

n_1 / n_2	2	3	4	5	6	7	8	9	10
3				**2**	**2**	**2**	**2**	**2**	**3**
4			**2**	**2**	**3**	**3**	**3**	**3**	**3**
5		**2**	**2**	**3**	**3**	**3**	**3**	**4**	**4**

6		**2**	**3**	**3**	**3**	**4**	**4**	**4**	**5**
7		**2**	**3**	**3**	**4**	**4**	**4**	**5**	**5**
8	**2**	**2**	**3**	**3**	**4**	**4**	**5**	**5**	**6**
9	**2**	**2**	**3**	**4**	**4**	**5**	**5**	**6**	**6**
10	**2**	**3**	**3**	**4**	**5**	**5**	**6**	**6**	**6**

Decyzje:

Jeśli $U_n \in K$ to H_0 odrzucamy ,

Jeśli $U_n \notin K$ to nie ma podstaw do odrzucenia H_0 .

Badanie symetrii składnika losowego.

Niech

n - liczba obserwacji,

m - liczba reszt dodatnich.

Wysuwamy dwie hipotezy:

$$H_0\left(\frac{m}{n}=\frac{1}{2}\right) \qquad\qquad H_1\left(\frac{m}{n}\neq\frac{1}{2}\right)$$

Stosujemy statystykę

$$U_n = \frac{\frac{m}{n}-\frac{1}{2}}{\sqrt{\frac{\frac{m}{n}\left(1-\frac{m}{n}\right)}{n-1}}}$$

Rozpatrujemy zbiór krytyczny:

$$K=(-\infty;-k>\cup<k;+\infty)$$

gdzie k odczytujemy dla poziomu istotności α z tablicy rozkładu Studenta:

$$P\left(\left|T_{n-1}\right|>k\right)=\alpha .$$

Decyzje:

Jeśli $U_n \in K$ to H_0 odrzucamy ,

Jeśli $U_n \notin K$ to nie ma podstaw do odrzucenia H_0.

(zmienność wyjaśniona)**Badanie normalności rozkładu reszt**

Zastosujemy **test Shapiro-Wilka**.

Wysuwamy dwie hipotezy:

H_0 - reszty mają rozkład normalny, H_1 - reszty nie mają rozkładu normalnego.

Reszty porządkujemy niemalejąco:

$$e_{(1)}, e_{(2)}, ..., e_{(n)}$$

Stosujemy statystykę

$$U_n = \frac{\left[\sum_{i=1}^{[n/2]} a_{n-i+1}\left(e_{(n-i+1)} - e_{(i)}\right)\right]^2}{\sum_{i=1}^{n}\left(e_i - \bar{e}\right)^2}$$

gdzie [n/2] jest częścią całkowitą liczby n/2,

$\bar{e}$ = 0 dla modeli liniowych,

a_{n-i+1} - współczynniki Shapiro-Wilka odczytane z tablicy:

i n	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
2	0,7071	-	-	-	-	-	-	-	-	-
3	0,7071	0	-	-	-	-	-	-	-	-
4	0,6	0,1	-	-	-	-	-	-	-	-

	872	677								
5	0,6 646	0,2 413	0	-	-	-	-	-	-	-
6	0,6 431	0,2 806	0,0 875	-	-	-	-	-	-	-
7	0,6 233	0,3 031	0,1 401	0	-	-	-	-	-	-
8	0,6 052	0,3 164	0,1 743	0,0 561	-	-	-	-	-	-
9	0,5 888	0,3 244	0,1 976	0,0 947	0	-	-	-	-	-
10	0,5 739	0,3 291	0,2 141	0,1 224	0,0 399	-	-	-	-	-
11	0,5 601	0,3 315	0,2 260	0,1 429	0,0 695	0	-	-	-	-
12	0,5 475	0,3 325	0,2 347	0,1 586	0,0 922	0,0 303	-	-	-	-
13	0,5 359	0,3 325	0,2 412	0,1 707	0,1 099	0,0 539	0	-	-	-
14	0,5 251	0,3 318	0,2 460	0,1 802	0,1 240	0,0 727	0,0 240	-	-	-
15	0,5 150	0,3 306	0,2 495	0,1 878	0,1 353	0,0 880	0,0 433	0	-	-
16	0,5 056	0,3 290	0,2 521	0,1 939	0,1 447	0,1 005	0,0 593	0,0 196	-	-
17	0,4 968	0,3 273	0,2 540	0,1 988	0,1 524	0,1 109	0,0 725	0,0 359	0	-

18	0,4886	0,3253	0,2553	0,2027	0,1587	0,1197	0,0837	0,0496	0,0163	-
19	0,4808	0,3232	0,2561	0,2059	0,1641	0,1271	0,0932	0,0612	0,0303	0
20	0,4734	0,3211	0,2565	0,2085	0,1686	0,1334	0,1013	0,0711	0,0422	0,0140

Rozpatrujemy zbiór krytyczny:

$$K =< 0;\ k >$$

gdzie k odczytujemy dla poziomu istotności α i danego n z tablicy testu Shapiro-Wilka:

(tablica testu Shapiro-Wilka dla $\alpha = 0{,}05$)

n	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
k	0,767	0,748	0,762	0,788	0,803	0,818	0,829	0,842	0,850	0,859

n	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
k	0,866	0,874	0,881	0,887	0,892	0,897	0,901	0,905

Decyzje:

Jeśli $U_n \in K$ to H_0 odrzucamy ,

Jeśli $U_n \notin K$ to nie ma podstaw do odrzucenia H_0 .

Badanie jednorodności wariancji składnika losowego

Jednorodność wariancji składnika losowego (homoskedastyczność) jest jednym z założeń klasycznej metody najmniejszych kwadratów. Niespełnienie tego założenia obniża efektywność estymatorów parametrów strukturalnych (nie wpływa na zgodność i nieobciążoność).

Zastosujemy **test Goldfelda-Quandta**.

W teście tym dzielimy próbę na dwie równoliczne podpróby o liczebnościach $n_1 = n_2$ (gdy liczba obserwacji jest nieparzysta - środkowa lub środkowe obserwacje nie biorą udziału w dalszych obliczeniach). Na podstawie tych podprób szacujemy parametry strukturalne modelu i obliczamy wariancje resztowe S_{e1}^2 , S_{e2}^2 . Próby numerujemy tak aby $S_{e2}^2 \geq S_{e1}^2$.

Wysuwamy dwie hipotezy:

$$H_0\left(\sigma_1^2 = \sigma_2^2\right) \qquad H_1\left(\sigma_2^2 > \sigma_1^2\right)$$

Stosujemy statystykę

$$U_n = \frac{S_{e2}^2}{S_{e1}^2}$$

Rozpatrujemy zbiór krytyczny:

$$K = < k; +\infty)$$

gdzie k odczytujemy dla poziomu istotności α z tablicy rozkładu Fiszera-Snedecora

dla $(n_2 - (k + 1), n_1 - (k + 1))$ stopni swobody.

Decyzje:

Jeśli $U_n \in K$ to H_0 odrzucamy ,

Jeśli $U_n \notin K$ to nie ma podstaw do odrzucenia H_0 .

Badanie autokorelacji reszt (test Durbina-Watsona).

Rozpatrujemy hipotezę: H_0(reszty nie są skorelowane) tzn $H_0(\rho = 0)$

Obliczamy wartość statystyki

$$U_n = \frac{\sum_{i=2}^{n} (e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^{\,2}}$$

Uwaga

1) $U_n \in < 0; 4 >$

2)Dla reszt nieskorelowanych $U_n \approx 2$

Z tablicy rozkładu D-W odczytuje się dla ustalonego α, k, n dwie liczby k_L i k_U .

Tablica rozkładu D-W dla $\alpha = 0{,}05$:

n	k = 1		k = 2	
	k_L	k_U	k_L	k_U
6	0,610	1,400	-	-
7	0,700	1,356	0,467	1,897
8	0,730	1,332	0,559	1,777
9	0,824	1,320	0,629	1,699
10	0,879	1,320	0,697	1,641
11	0,927	1,324	0,758	1,604
12	0,971	1,331	0,812	1,579
13	1,010	1,340	0,861	1,562
14	1,045	1,350	0,905	1,552
15	1,077	1,361	0,946	1,543

Jeśli $U_n < 2$ to rozpatrujemy hipotezę alternatywną:
H_1(reszty są skorelowane dodatnio) tzn $H_1(\rho > 0)$.

Przyjmuje się następującą regułę decyzyjną:

Jeśli $U_n < \mathrm{k_L}$ to H_0 odrzucamy.

Jeśli $U_n > \mathrm{k_U}$ to nie ma podstaw do odrzucenia H_0.

Jeśli $k_L \le U_n \le k_U$ to nie podejmujemy decyzji.

Jeśli $U_n > 2$ to rozpatrujemy hipotezę alternatywną:

H_1(reszty są skorelowane ujemnie) tzn $H_1(\rho < 0)$.

Przyjmuje się następującą regułę decyzyjną:

Jeśli $U_n > 4 - \mathrm{k_L}$ to H_0 odrzucamy.

Jeśli $U_n < 4 - \mathrm{k_U}$ to nie ma podstaw do odrzucenia H_0.

Jeśli $4 - \mathrm{k_U} \le U_n \le 4 - \mathrm{k_L}$ to nie podejmujemy decyzji.

Prognoza.

Niech $x_\tau = [1 \quad x_{\tau 1} \quad \cdots \quad x_{\tau k}]^T$

Prognoza punktowa

$$\boxed{y_\tau = x_\tau^T b}$$

Standardowy błąd prognozy

$$\boxed{S_\tau = S_e \sqrt{1 + x_\tau^T \left(X^T X \right)^{-1} x_\tau}}$$

gdy k = 1

$$\boxed{S_\tau = S_e \sqrt{1 + \frac{1}{n} + \frac{(x_\tau - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}} = S_e \sqrt{1 + \frac{\sum_{i=1}^{n} x_i^{\,2} + n x_\tau^2 - 2 x_\tau \sum_{i=1}^{n} x_i}{n \sum_{i=1}^{n} x_i^{\,2} - \left(\sum_{i=1}^{n} x_i \right)^2}}}$$

Prognoza przedziałowa dla poziomu ufności 1-α.

$$\boxed{\langle y_\tau - u_\alpha S_\tau ; y_\tau + u_\alpha S_\tau \rangle}$$

gdzie u_α odczytujemy z tablicy rozkładu Studenta: $P\left(\left|T_{n-(k+1)}\right| > u_\alpha\right) = \alpha$.

Przykład.

Y - wielkość produkcji (tys. szt.),

X_1 - liczba zatrudnionych (tys. osób),

X_2 - wartość majątku trwałego (mln zł),

Rozpatrujemy model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$. Mając dane

Rok	1992	1993	1994	1995	1996	1997	1998
Y	76,9	90,2	95,5	100	102,4	107	110,5
X_1	4,5	4,7	4,8	4,8	5,0	5,2	5,0
X_2	11	16,5	17	17,2	18,4	20	21,6

oraz wiedząc, że

$$\left(X^T X\right)^{-1} = \begin{bmatrix} 216{,}560678 & -55{,}768887 & 3{,}132436 \\ -55{,}768887 & 14{,}675376 & -0{,}892199 \\ 3{,}132436 & -0{,}892199 & 0{,}069085 \end{bmatrix}$$

$$\hat{Y} = -1{,}1862 + 10{,}8311 X_1 + 2{,}6503 X_2$$

oraz $S_e = 2{,}53$

wyznaczymy prognozę punktową i przedziałową na rok 1999 dla

$$x_\tau = \begin{bmatrix} 1 & 7 & 25 \end{bmatrix}^T .$$

Rozwiązanie:

Ponieważ

$$b = \begin{bmatrix} -1{,}1862 \\ 10{,}8311 \\ 2{,}6503 \end{bmatrix};$$

to wartość prognozy punktowej jest równa:

$$y_\tau = x_\tau^T b = \begin{bmatrix} 1 & 7 & 25 \end{bmatrix} \begin{bmatrix} -1{,}1862 \\ 10{,}8311 \\ 2{,}6503 \end{bmatrix} = 140{,}89$$

Ponieważ

$$x_\tau^T \left(X^T X \right)^{-1} x_\tau = \begin{bmatrix} 1 & 7 & 25 \end{bmatrix} \begin{bmatrix} 216{,}560678 & -55{,}768887 & 3{,}132436 \\ -55{,}768887 & 14{,}675376 & -0{,}892199 \\ 3{,}132436 & -0{,}892199 & 0{,}069085 \end{bmatrix} \begin{bmatrix} 1 \\ 7 \\ 25 \end{bmatrix} = 42{,}5$$

to standardowy błąd prognozy wynosi

$$S_\tau = S_e \sqrt{1 + x_\tau^T \left(X^T X \right)^{-1} x_\tau} = 2{,}53 \sqrt{1 + 42{,}5} = 16{,}67$$

Zatem przewidywana wielkość produkcji wynosi $140 \pm 16{,}67$.

Prognoza przedziałowa dla poziomu ufności $1-\alpha = 0{,}95$.

Liczbę u_α odczytujemy z tablicy rozkładu Studenta: $P(|T_4| > u_\alpha) = 0{,}05$ u_α = 2,78.

$$\langle y_\tau - u_\alpha S_\tau ; y_\tau + u_\alpha S_\tau \rangle = <94{,}61;\ 187{,}17>$$

DODATEK 1.

Uogólniona metoda najmniejszych kwadratów.

Jeśli wariancja składników losowych nie jest stała (brak homoscedastyczności) lub nie są spełnione założenia o braku autokorelacji reszt to należy do szacowania parametrów strukturalnych stosować uogólnioną metodę najmniejszych kwadratów:

$$b = \left(X^T V^{-1} X\right)^{-1} X^T V^{-1} Y$$

gdzie V - dodatnio określona macierz symetryczna stopnia n.

W przypadku braku homoscedastyczności można np. przyjąć:

$$V = \begin{bmatrix} |e_1| & 0 & \cdots & 0 \\ 0 & |e_2| & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & |e_n| \end{bmatrix}$$

gdzie e_i - reszty modelu oszacowane MNK.

W przypadku autokorelacji reszt można np. przyjąć:

$$V = \begin{bmatrix} 1 & r & \cdots & r^{n-1} \\ r & 1 & \cdots & r^{n-2} \\ \cdots & \cdots & \cdots & \cdots \\ r^{n-1} & r^{n-2} & \cdots & 1 \end{bmatrix}$$

gdzie

$$r = \frac{(n-k-1)\sum_{i=1}^{n-1} e_i e_{i+1}}{(n-1)\sum_{i=1}^{n} e_i^2}$$

e_i - reszty modelu oszacowane MNK.

DODATEK 2.

Dobór zmiennych objaśniających (model liniowy).

$X_1, X_2,..., X_k$ – zmienne objaśniające, Y - zmienna objaśniana,

Zmienne objaśniające powinny charakteryzować się:

a) wysoką zmiennością (współczynnik zmienności powyżej określonej wartości krytycznej np., $V(X_i) > 0{,}1$),
b) silną korelacją z Y,
c) słabą korelacją z innymi zmiennymi objaśniającymi.

Przykład.

Mając dane wartości zmiennych

Y	X_1	X_2	X_3	X_4
5	3	10	7	6
5	4	8	7	5
8	5	7	6	7
8	6	7	7	7
6	5	6	9	6
7	5	5	10	6
10	7	5	12	6
10	7	4	10	7
12	6	4	11	6
12	8	4	12	6

Sprawdź, które zmienne należy wyeliminować jako quasi stałe przyjmując krytyczną wartość współczynnika zmienności równą 0,15?
Współczynniki zmienności dla poszczególnych zmiennych objaśniających są równe:

$V(X_1)$	$V(X_2)$	$V(X_3)$	$V(X_4)$
0,255	0,316	0,233	0,097

Zatem należy wyeliminować zmienną X_4 .

Jeśli zmienne X, Y mają pary wartości (x_i, y_i) to współczynnik korelacji Pearsona obliczamy następująco:

$$r_{XY} = \frac{\operatorname{cov}(X,Y)}{S_X \cdot S_Y} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sqrt{\sum x_i^2 - n(\bar{x})^2} \sqrt{\sum y_i^2 - n(\bar{y})^2}}$$

Niech $r_i = r_{X_i Y}$ - współczynniki korelacji między poszczególnymi zmiennymi objaśniającymi a zmienną objaśnianą. Wektorem korelacji nazywamy wektor

$$R_0 = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_k \end{bmatrix}$$

Niech $r_{ij} = r_{X_i X_j}$ - współczynniki korelacji między poszczególnymi zmiennymi objaśniającymi. Macierzą korelacji nazywamy symetryczną macierz

$$R = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1k} \\ r_{21} & 1 & \cdots & r_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ r_{k1} & r_{k2} & \cdots & 1 \end{bmatrix}$$

Metoda wyboru zmiennych objaśniających, które są silnie skorelowane ze zmienną objaśniającą i słabo skorelowane z innymi zmiennymi objaśniającymi.

Kolejność postępowania:

1) ustalamy krytyczną wartość korelacji r_{kryt} (albo jest narzucona albo obliczamy ze wzoru

$r_{kryt} = \sqrt{\frac{k_\alpha^2}{k_\alpha^2 + n - 2}}$ gdzie k_α - odczytujemy z tablicy rozkładu Studenta dla n - 2 stopni swobody i poziomu istotności α.

2) eliminujemy te zmienne objaśniające dla których:

$$|r_i| \leq r_{kryt}$$

3) spośród pozostałych zmiennych **wybieramy** taką zmienną X_s dla której

$|r_s| = \max \{|r_i|\}$ (ta zmienna niesie najwięcej informacji)

4) ze zbioru zmiennych objaśniających **eliminujemy** te dla których

$$|r_{si}| > r_{kryt}$$

(zmienne silnie skorelowane z wybraną zmienną X_s powielają zawarte w X_s informacje).

Kroki 3) i 4) można ewentualnie powtarzać.

Przykład.

Dla zmiennych X_1, X_2, X_3, X_4 i Y z poprzedniego przykładu wektor korelacji i macierz korelacji są równe:

$$R_0 = \begin{bmatrix} 0,88 \\ -0,82 \\ 0,73 \\ 0,29 \end{bmatrix} \qquad R = \begin{bmatrix} 1 & -0,85 & 0,74 & 0,33 \\ -0,85 & 1 & -0,82 & -0,18 \\ 0,74 & -0,82 & 1 & -0,17 \\ 0,33 & -0,18 & -0,17 & 1 \end{bmatrix}$$

1) Dla poziomu istotności 0,05 i 10 - 2 = 8 stopni swobody odczytujemy z tablicy rozkładu Studenta $k_\alpha = 2,306$ i wyznaczamy

$$r_{kryt} = \sqrt{\frac{k_\alpha^2}{k_\alpha^2 + n - 2}} = \sqrt{\frac{2,306^2}{2,306^2 + 10 - 2}} = 0,63$$

2) odrzucamy zmienną X_4,

Zredukowany wektor i zredukowana macierz korelacji są równe

$$R_0 = \begin{bmatrix} 0,88 \\ -0,82 \\ 0,73 \end{bmatrix} \qquad R = \begin{bmatrix} 1 & -0,85 & 0,74 \\ -0,85 & 1 & -0,82 \\ 0,74 & -0,82 & 1 \end{bmatrix}$$

3) wybieramy X_1 ,

4) eliminujemy X_2, X_3,

Zatem rozpatrywany model liniowy powinien mieć postać:

$$Y = \beta_0 + \beta_0 X_1 + \varepsilon$$

Metoda Hellwiga.

Rozpatrujemy wszystkie niepuste podzbiory zbioru zmiennych objaśniających

$\{X_1, X_2, \ldots, X_k\}$, takich podzbiorów jest $L = 2^k - 1$.

Dla każdego podzbioru oblicza się wskaźniki pojemności informacyjnej: indywidualne i

integralne (ich wartości należą do przedziału [0, 1]).

Indywidualną pojemność informacyjną obliczamy ze wzoru:

$$h_{lj} = \frac{r_j^2}{\sum_{i \in I_l} |r_{ij}|}$$

gdzie l = 1, 2, ..., L (numer podzbioru - kombinacji), I_l - zbiór numerów zmiennych wchodzących w skład l - tego podzbioru.

Integralną pojemność informacyjną obliczamy sumując pojemności indywidualne rozpatrywanego podzbioru:

$$H_l = \sum_{j \in I_l} h_{lj}$$

Należy wybrać taki podzbiór zmiennych objaśniających dla którego integralna pojemność informacyjna jest maksymalna.

Przykład.

Dla zmiennych X_1, X_2, Y obliczono

$$R_0 = \begin{bmatrix} 0{,}8 \\ -0{,}2 \end{bmatrix} \qquad R = \begin{bmatrix} 1 & 0{,}6 \\ 0{,}6 & 1 \end{bmatrix}$$

Są 3 podzbiory zbioru $\{X_1, X_2\}$: $\{X_1\}$, $\{X_2\}$, $\{X_1, X_2\}$.

Obliczamy:

$h_{11} = 0{,}8^2 = 0{,}64$, $H_1 = 0{,}64$,

$h_{22} = (-0,2)^2 = 0,04$, $H_2 = 0,04$,
$h_{31} = 0,8^2/(1 + 0,6) = 0,4$, $h_{32} = (-0,2)^2/(1 + 0,6) = 0,025$, $H_3 = h_{31} + h_{32} = 0,425$,
Ponieważ największą pojemność informacyjną ma podzbiór $\{X_1\}$, to należy przyjąć, że X_1 jest jedyną zmienną objaśniającą w tym modelu tzn. $Y = \beta_0 + \beta_0 X_1 + \varepsilon$.

DODATEK 3.

Modele liniowe wielorównaniowe.

Zjawiskom ekonomicznym wyjaśnianym przez model wielorównaniowy odpowiadają zmienne **endogeniczne**.
Pozostałe zmienne nazywamy zmiennymi **egzogenicznymi**.
Y_1, ..., Y_m - zmienne endogeniczne bez opóźnień czasowych
Z_1, ..., Z_k - zmienne endogeniczne z opóźnieniami czasowymi i zmienne egzogeniczne
Ogólny zapis modelu:

$$Y_1 = \sum_{i=2}^{m} \beta_{1i} Y_i + \sum_{j=1}^{k} \gamma_{1j} Z_j + \varepsilon_1$$

$$Y_2 = \sum_{\substack{i=1 \\ i \neq 2}}^{m} \beta_{2i} Y_i + \sum_{j=1}^{k} \gamma_{2j} Z_j + \varepsilon_{21}$$

.....

$$Y_m = \sum_{i=1}^{m-1} \beta_{mi} Y_i + \sum_{j=1}^{k} \gamma_{mj} Z_j + \varepsilon_m$$

Macierzowe przedstawienie tego zapisu nosi nazwę postaci strukturalnej:

$$BY + \Gamma Z = \varepsilon$$

gdzie

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

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_m \end{bmatrix} \qquad Z = \begin{bmatrix} Z_1 \\ Z_2 \\ \dots \\ Z_k \end{bmatrix} \qquad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_m \end{bmatrix}$$

Jeśli zmienne Y_1, ..., Y_m wyrazimy przez Z_1, ..., Z_k to otrzymamy postać zredukowaną modelu:

$$Y_1 = \sum_{j=1}^{k} \pi_{1j} Z_j + \eta_1$$

$$Y_2 = \sum_{j=1}^{k} \pi_{2j} Z_j + \eta_2$$

.....

$$Y_m = \sum_{j=1}^{k} \pi_{mj} Z_j + \eta_m$$

Postać macierzowa:

$$Y = \Pi^T Z + \eta$$

gdzie

$$\Pi^T = \begin{bmatrix} \pi_{11} & \pi_{12} & \dots & \pi_{1k} \\ \pi_{21} & \pi_{22} & \dots & \pi_{2k} \\ \dots & \dots & \dots & \dots \\ \pi_{m1} & \pi_{m2} & \dots & \pi_{mk} \end{bmatrix} \qquad \eta = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \dots \\ \eta_m \end{bmatrix}$$

Powyższą macierz i wektor wyznaczamy ze wzorów:

$$\Pi^T = -B^{-1}\Gamma \qquad \eta = B^{-1}\varepsilon$$

Klasyfikacja modeli wielorównaniowych

1. Jeśli macierz B jest macierzą diagonalną (ewentualnie po przenumerowaniu równań) to model nazywamy **prostym**,
2. Jeśli macierz B jest macierzą trójkątną (ewentualnie po przenumerowaniu równań lub zmiennych) to model nazywamy **rekurencyjnym**,
3. W pozostałych przypadkach model nazywamy modelem **o równaniach współzależnych**.

Parametry modeli prostych i rekurencyjnych szacujemy jak parametry modeli jednorównaniowych (każde równanie możemy rozpatrywać oddzielnie).

Parametry modelu o zmiennych współzależnych można oszacować tylko wtedy gdy wszystkie jego równania są **idetyfikowalne.**

Twierdzenie.
Warunkiem koniecznym i dostatecznym tego, aby i - te równanie modelu o m równaniach współzależnych było identyfikowalne, jest macierz A_i parametrów znajdujących się przy zmiennych, które są w modelu, a nie występują w równaniu,

którego identyfikowalność jest badana, była rzędu m - 1.

Niech k_i - liczba zmiennych, znajdujących się w modelu, a nie występują w równaniu, którego identyfikowalność jest badana.
Jeśli $k_i = m - 1$, to mówimy, że równanie jest **jednoznacznie identyfikowalne**.
Jeśli $k_i > m - 1$, to mówimy, że równanie jest niejednoznacznie identyfikowalne.
Jeśli $k_i < m - 1$, to mówimy, że równanie nie jest identyfikowalne.

Parametry modelu o zmiennych współzależnych i równaniach jednoznacznie identyfikowalnych można oszacować metodą najmniejszych kwadratów:

$$P = \left(Z^T Z\right)^{-1} Z^T Y \qquad \Gamma = -BP^T$$

DODATEK 4.
REGRESJA KRZYWOLINIOWA.

Parametry wybranej funkcji nieliniowej wyznacza się metodą najmniejszych kwadratów, korzystając z odpowiedniego układu równań normalnych.

Funkcja wielomianowa.

$$\hat{y}_i = b_0 + b_1 x + b_2 x^2 + ... + b_k x^k$$

Jej parametry $b_0, ..., b_k$ wyznaczamy rozwiązując układ równań normalnych który ma postać:

$$\begin{cases} nb_0 + \left(\sum_i x_i\right) b_1 + \left(\sum_i x_i^2\right) b_2 + ... + \left(\sum_i x_i^k\right) b_k = \sum_i y_i \\ \left(\sum_i x_i\right) b_0 + \left(\sum_i x_i^2\right) b_1 + \left(\sum_i x_i^3\right) b_2 + ... + \left(\sum_i x_i^{k+1}\right) b_k = \sum_i x_i y_i \\ ... \\ \left(\sum_i x_i^k\right) b_0 + \left(\sum_i x_i^{k+1}\right) b_1 + \left(\sum_i x_i^{k+2}\right) b_2 + ... + \left(\sum_i x_i^{2k}\right) b_k = \sum_i x_i^k y_i \end{cases}$$

Powyższy układ równań otrzymujemy przyrównując do zera pochodne cząstkowe funkcji k + 1 zmiennych

$$S(b_0, b_1, b_2, ..., b_k) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} \left(y_i - \left(b_0 + b_1 x_i + b_2 x_i^2 + ... + b_k x_i^k\right)\right)^2$$

W szczególności **Funkcja kwadratowa.**

$$\hat{y}_i = b_0 + b_1 x + b_2 x^2$$

Jej parametry $b_0, ..., b_2$ wyznaczamy rozwiązując układ równań normalnych który ma postać:

$$\begin{cases} nb_0 + \left(\sum_i x_i\right) b_1 + \left(\sum_i x_i^2\right) b_2 = \sum_i y_i \\ \left(\sum_i x_i\right) b_0 + \left(\sum_i x_i^2\right) b_1 + \left(\sum_i x_i^3\right) b_2 = \sum_i x_i y_i \\ \left(\sum_i x_i^2\right) b_0 + \left(\sum_i x_i^3\right) b_1 + \left(\sum_i x_i^4\right) b_2 = \sum_i x_i^2 y_i \end{cases}$$

Funkcja potęgowa.

$$\hat{y}_i = a x_i^{\,b}$$

Chociaż jest to szczególny przypadek funkcji wielomianowej to warto rozpatrywać go również oddzielnie.

Jej parametry a, b wyznaczamy przez przekształcenie do postaci liniowej (logarytmujemy obie strony).

$$\ln \hat{y}_i = \ln a + b \ln x_i$$

Układ równań normalnych ma postać:

$$\begin{cases} \sum_i \ln y_i = n \ln a + b \sum_i \ln x_i \\ \sum_i \ln x_i \ln y_i = \ln a \sum_i \ln x_i + b \sum_i \ln^2 x_i \end{cases}$$

Rozwiązując powyższy układ równań (liniowych względem a' = lna i b) obliczamy a' i b. Stąd a = $e^{a'}$.

Parametr b jest interpretowany jako współczynnik elastyczności, tzn. jeśli zmienna X wzrośnie o 1%, to Y zmieni się średnio o b %.

Funkcja wykładnicza.

$$\hat{y}_i = ab^{x_i}$$

Logarytmując obie strony otrzymamy.

$$\ln \hat{y}_i = \ln a + x_i \ln b$$

Układ równań normalnych ma postać:

$$\begin{cases} \sum_i \ln y_i = n \ln a + \ln b \sum_i x_i \\ \sum_i x_i \ln y_i = \ln a \sum_i x_i + \ln b \sum_i x_i^2 \end{cases}$$

Rozwiązując powyższy układ równań (liniowych względem a' = lna i b' = lnb) obliczamy a' i b'. Stąd $a = e^{a'}$ i $b = e^{b'}$.

Parametr b jest interpretowany jako średni przyrost względny, tzn. jeśli zmienna X wzrośnie o jednostkę, to Y zmieni się średnio o (b - 1)100 %.

Funkcja logistyczna.

$$\hat{y}_i = \frac{a}{1 + be^{-ct}}$$

gdzie t - czas, a > 0, b > 1, c > 0.

Funkcja logistyczna służy między innymi do opisu i prognozowania ,wielkości sprzedaży produktu wchodzącego na rynek.
Przyjmujemy

$$z_t = \frac{1}{y_t}$$

Najpierw wyznaczamy wartości parametrów a, c

$$a = \frac{u}{u_0} - 1 \qquad\qquad c = -\ln\frac{u_1}{u}$$

gdzie

$$u = (n-1)\sum_{t=1}^{n-1} z_t^2 - \left(\sum_{t=1}^{n-1} z_t\right)^2$$

$$u_0 = \sum_{t=1}^{n-1} z_{t+1} \sum_{t=1}^{n-1} z_t^2 - \sum_{t=1}^{n-1} z_t \sum_{t=1}^{n-1} z_t\, z_{t+1}$$

$$u_1 = (n-1)\sum_{t=1}^{n-1} z_t\, z_{t+1} - \sum_{t=1}^{n-1} z_{t+1} \sum_{t=1}^{n-1} z_t$$

Następnie korzystając z obliczonych a i c obliczamy b

$$b = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{a}{y_t} - 1\right)e^{ct} \qquad t = 1, 2, ...,n$$

Parametr b > 1 gwarantuje istnienie punktu przegięcia, a jest interpretowany jako poziom nasycenia (asymptota pozioma).

W przypadku nieliniowym miarą dopasowania dopasowania modelu do danych statystycznych jest współczynnik korelacji krzywoliniowej

$$R = \sqrt{1 - \frac{\sum(y_i - \hat{y})^2}{\sum(y_i - \bar{y})^2}}; \qquad R \in \langle 0, 1\rangle.$$

Przykład.

Mając dane

t	1	2	3	4	5	6	7
y	8	6	5	2	2	4	7

Wyznaczymy kwadratową funkcję regresji.

Korzystając z sum w poniższej tabeli układamy układ równań normalnych:

t	y	t^2	t^3	t^4	y^2	ty	t^2y
1	8	1	1	1	64	8	8
2	6	4	8	16	36	12	24
3	5	9	27	81	2	1	45

					5	5	
4	2	16	64	256	4	8	32
5	2	25	125	625	4	10	50
6	4	36	216	1296	16	24	144
7	7	49	343	2401	49	49	343
8	8	64	512	4096	64	64	512
suma	36	42	204	1296	8772	2620	1190	1158

$a \cdot 8772 + b \cdot 1296 + c \cdot 204 = 1158$

$a \cdot 1296 + b \cdot 204 + c \cdot 36 = 190$

$a \cdot 204 + b \cdot 36 + c \cdot 8 = 42$

Rozwiązaniem (przybliżonym) układu jest

$a = 0{,}4643; \quad b = -4{,}1548; \quad c = 12{,}107$

Zatem funkcja regresji kwadratowej ma postać

$$\hat{y} = 0{,}4643x^2 - 4{,}1548x + 12{,}107$$

Funkcja ta jest dobrze dopasowana do danych statystycznych ($R^2 = 0{,}8732$, $R = 0{,}93$).

Zauważmy, że w tym przypadku funkcja liniowa nie jest dobrą funkcją regresji a bardzo niska wartość współczynnika korelacji

Pearsona świadczy o braku zależności liniowej a nie o braku zależności jakiejkolwiek.