Ekonometria od podstaw

Joanna Głowacka

www.bombowaksiegowa.pl

Wrocław, 2017

Spis treści

1. Co to jest ekonometria i model ekonometryczny?

Ekonometria to nauka zajmująca się modelowaniem procesów ekonomicznych. Inaczej mówiąc ilościowo (liczbowo) opisuje prawidłowości ekonomiczne w oparciu o dane statystyczne. Prawidłowości te zostały opisane przez teorię ekonomii lub sugerowane przez hipotezy ekonomiczne. Ekonometria wykorzystuje tzw. **model ekonometryczny**. Jest to model umożliwiający opisanie danego zjawiska ekonomicznego. Istotne jest to, że aby zapisać ten model musimy zrozumieć jego mechanizm. Wyróżniamy 3 cele budowy modelu:

- cel poznawczo-opisowy -> chcemy opisać mechanizm kształtowania się zjawisk ekonomicznych,
- cel diagnostyczno-kontrolny -> chcemy ocenić zbadany mechanizm,
- cel predyktywny -> chcemy przewidzieć przebieg zjawisk ekonomicznych.

W modelu ekonometrycznym istnieją dwa rodzaje zmiennych:

- Y- zmienna objaśniana, zależna
- X- zmienna objaśniająca, niezależna

Jak to zapamiętać? X wyjaśnia Y, dlatego jest objaśniająca i niezależna, bo nie zależy od innej zmiennej. Natomiast Y jest zależna od X i objaśnia ją X. Wyróżnia się dwa rodzaje modeli:

- model z jedną zmienną objaśniającą
- model regresji wielorakiej

Model z jedną zmienną objaśniająca

Jak sama nazwa wskazuje zawiera tylko jednego X i oczywiście Y. Jego postać jest następująca:

$$Y = \alpha_0 + \alpha_1 * X + \varepsilon$$

Wyjaśnienie pozostałych zmiennych:

- α_0- wyraz wolny
- α_1- współczynnik kierunkowy
- ε- posiada różne nazwy: zaburzenia losowe, zakłócenia, błąd losowy, składnik losowy. O nim dowiemy się więcej w dalszych rozdziałach.

Model regresji wielorakiej

Model ten zawiera więcej niż 1 zmiennych objaśniających. Ich liczba może być nieskończona, o ile zmienne te wpływają na Y.

$$Y = \alpha_0 + \alpha_1 * X_1 + \alpha_2 * X_2 + \alpha_3 * X_3 + \cdots + \alpha_n * X_n + \varepsilon$$

Etapy budowy modelu ekonometrycznego

W przypadku tworzenia jednego modelu wyróżniamy kilka etapów:

1. Wybór zmiennych do modelu
2. Sprawdzenie istotności korelacji
3. Oszacowanie parametrów modelu
4. Ocena istotności parametrów
5. Interpretacja parametrów i ich zgodność z teorią ekonomii
6. Ocena specyfikacji modelu oraz transformacje
7. Weryfikacja założeń MNK

2. Wybór zmiennych do modelu

Mamy 3 możliwości wyboru zmiennych do modelu:

- Metoda eliminacji krokowej zmiennych- jest to metoda, która opiera się na wartości p i istotności parametrów (rozdział 5),
- Metoda współczynników korelacji,
- Metoda kryterium Akaike'a.

W metodzie współczynników korelacji tej ustala się tzw. wartość krytyczną współczynnika korelacji r^* wyznaczona według wzoru:

$$r^* = \min_i \max_j r_{ij}$$

lub

$$r^* = \sqrt{\frac{t_{n-2}^2}{n-2+t_{n-2}^2}}$$

Wyróżniamy 3 etapy tej metody:

1. Ze zbioru potencjalnych zmiennych objaśniających (wszystkie X) eliminuje się wszystkie zmienne za słabo (nieistotnie) skorelowane ze zmienną objaśnianą. Musimy stworzyć macierz korelacji między Y, a wszystkimi X. Jeżeli którakolwiek z tych wartości jest mniejsza niż r^* to eliminujemy tą zmienną z modelu.
2. Spośród pozostałych potencjalnych zmiennych jako zmienną objaśniającą wybiera się zmienną, która jest najsilniej skorelowana ze zmienną objaśnianą. Patrzymy na tą samą macierz i wybieramy największą wartość. Ta zmienna wchodzi do modelu.
3. Ze zbioru pozostałych potencjalnych zmiennych objaśniających eliminuje się wszystkie zmienne powielające informacje. Musimy stworzyć nową macierz korelacji tym razem między zmienną wybraną w pkt 2 (najsilniej skorelowaną z Y) a pozostałymi potencjalnymi zmiennymi objaśniającymi. Jeżeli któraś z wartości w tej macierzy przekracza r^* to eliminujemy tą zmienną z modelu, ponieważ jest to zmienna, która dostarcza te same informacje, co zmienna najsilniej skorelowana z Y.
4. Punkty 2 i 3 powtarzamy do momentu wyczerpania zmiennych objaśniających.

Najczęściej wykorzystywaną metodą jest metoda eliminacji krokowej, ponieważ jest najszybsza. W drugiej kolejności wykorzystywana jest metoda współczynników korelacji. Stosowana gdy liczba potencjalnych X jest większa od wartości obserwacji. Natomiast najrzadziej spotykaną metodą jest metoda kryterium Akaike'a.

3. Istotność korelacji

Korelacja to współzależność dwóch zmiennych. W statystyce i ekonometrii oznacza związek między zmiennymi X i Y. Oznacza to nic innego jak sprawdzenie czy zmienna X wpływa na zmienną Y i w jaki sposób. Wyróżniamy dwa główne współczynniki korelacji:

- Pearsona
- Rang Spearmana

Współczynnik korelacji Pearsona wykazuje zależność liniową między zmiennymi. Współczynnik oblicza się ze wzoru:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

Współczynnik korelacji przyjmuje wartości w przedziale: $< -1; 1 >$. Wartość świadczy o sile korelacji. Są pewne przedziały umożliwiające ocenę siły korelacji.

0,0 $brak\ korelacji$

$(0{,}0\ ;\ 0{,}2 >$ $bardzo\ słaba\ korelacja$

$(0{,}2\ ;\ 0{,}4 >$ $słaba\ korelacja$

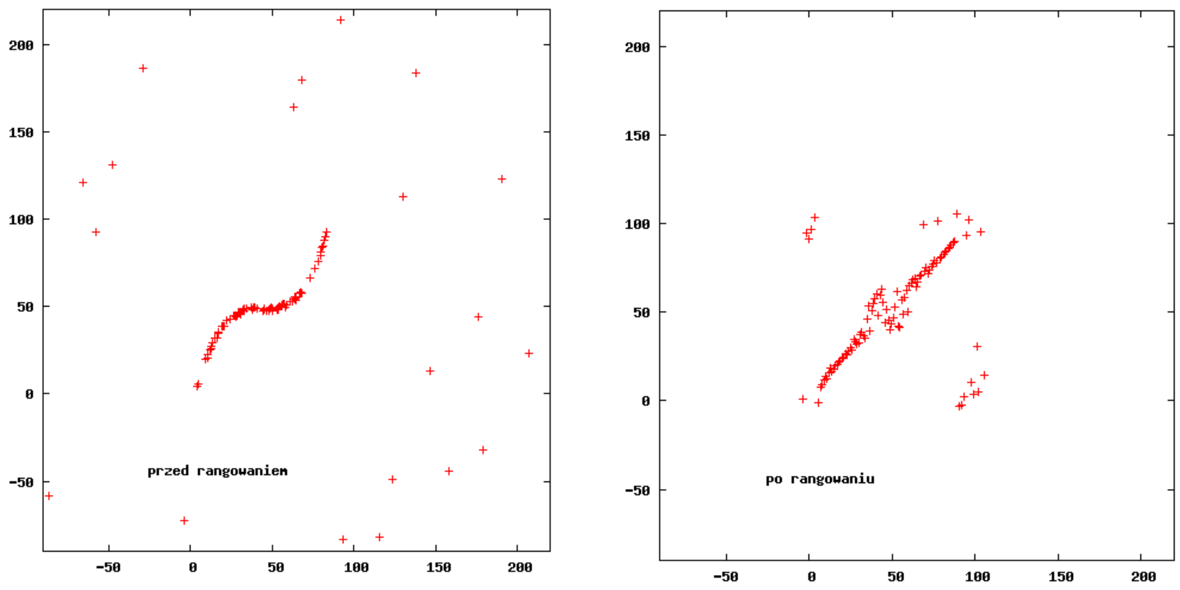
$(0{,}4\ ;\ 0{,}6 >$ $umiarkowana\ korelacja$

$(0{,}6\ ;\ 0{,}8 >$ $silna\ korelacja$

$(0{,}8\ ;\ 1)$ $bardzo\ silna\ korelacja$

1,0 $pełna\ korelacja$

Współczynnik korelacji rang Spearmana- Spearman zauważył, że w niektórych badaniach nie da się zastosować współczynnika Pearsona. Ten współczynnik umożliwia liczenie zwykłego współczynnika korelacji dla rang zmiennych. Obecnie występuje kilka wersji tego współczynnika. Poniższe wykresy wskazują na sposób jego działania:



Źródło: https://pl.wikipedia.org/wiki/Wsp%C3%B3%C5%82czynnik_korelacji_rang_Spearmana#/media/File:Spearman_animacja2.gif

Aby sprawdzić korelację dwóch zmiennych możemy zastosować test istotności korelacji:

$$H_0: \rho(X,Y) = 0$$

$$H_1: \rho(X,Y) \neq 0$$

Statystyka testowa:

$$T = \frac{|r| * \sqrt{n-2}}{\sqrt{1-r^2}}$$

W przypadku hipotez statystycznych zawsze dążymy do odrzucenia H_0.

$$H_0 \; odrzucamy \; gdy \; T > t_{\frac{\alpha}{2}, n-2}$$

Pamiętajcie o tym, że każdy test statystyczny musi składać się z kilku elementów: hipotezy, statystyka testowa, formułka o odrzuceniu H_0 i wniosek końcowy stwierdzający czy przyjmujemy/odrzucamy H_0 oraz co to znaczy dla nas.

Przykład:

Mamy dwie zmienne: dochody budżetowe na 1 mieszkańca oraz wydatki budżetowe na 1 mieszkańca. Mając następujące dane, zbadaj istotność korelacji między danymi:

```
corr(dochod, wydatki) = 0,97925083
Hipoteza zerowa: R = 0, brak korelacji:
 t(378) = 93,9483, przy dwustronym obszarze krytycznym p = 0,0000
```

```
Model 1: Estymacja KMNK, wykorzystane obserwacje 1-380
Zmienna zależna (Y): wydatki

             współczynnik  błąd standardowy  t-Studenta  wartość p
  ---------------------------------------------------------------
  const      -36,7302       17,1636          -2,140     0,0330    **
  dochod       1,04436       0,0111163       93,95     3,75e-264 ***

Średn.aryt.zm.zależnej  1532,046   Odch.stand.zm.zależnej  381,3449
Suma kwadratów reszt     2263480   Błąd standardowy reszt  77,38244
Wsp. determ. R-kwadrat  0,958932   Skorygowany R-kwadrat   0,958824
F(1, 378)               8826,288   Wartość p dla testu F   3,7e-264
Logarytm wiarygodności -2190,723   Kryt. inform. Akaike'a  4385,445
Kryt. bayes. Schwarza   4393,326   Kryt. Hannana-Quinna    4388,572
```

Rozwiązanie

Są dwa sposoby rozwiązania tego zadania. Oto pierwszy z nich:

1) Hipotezy:

$$H_0: \rho(dochód, wydatki) = 0$$

$$H_1: \rho(dochód, wydatki) \neq 0$$

2) Statystyka testowa:

$r = 0{,}9793$

$$T = \frac{|r| * \sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0{,}9793 * \sqrt{378}}{\sqrt{1-0{,}9793^2}} = 94{,}0634$$

$$H_0\ odrzucamy\ gdy\ T > t_{\frac{\alpha}{2},n-2}$$

Standardowo przyjmujemy $\alpha = 0{,}05$.

$t_{\frac{\alpha}{2},n-2} = t_{\frac{0{,}05}{2},380-2} = t_{0{,}025,378} = 1{,}966$

$$94{,}0634 > 1{,}966$$

3) Wniosek:

Odrzucamy H_0. Istnieje zależność liniowa między dochodami budżetowymi na 1 mieszkańca a wydatkami budżetowymi na 1 mieszkańca.

Drugi sposób na rozwiązanie tego zadania:

1) Hipotezy:

$$H_0: \rho(dochód, wydatki) = 0$$

$$H_1: \rho(dochód, wydatki) \neq 0$$

2) Zauważamy, że mamy podany dwustronny obszar krytyczny.

$$H_0 \; odrzucamy \; gdy \; p < \alpha$$

$$p = 0{,}0000 < 0{,}05$$

3) Wniosek:

   Odrzucamy H_0. Istnieje zależność liniowa między dochodami budżetowymi na 1 mieszkańca a wydatkami budżetowymi na 1 mieszkańca

4. Oszacowanie parametrów modelu

Po co szacowanie parametrów modelu? Nie da się zbadać wszystkich czynników wpływających na dane zjawisko. Z całej populacji wybieramy konkretną próbę reprezentatywną, tzn. taką, dzięki której nasz model może odzwierciedlać dane zjawisko. Oszacowany model wygląda następująco:

$$Y = \hat{\alpha}_o + \hat{\alpha}_1 * X$$

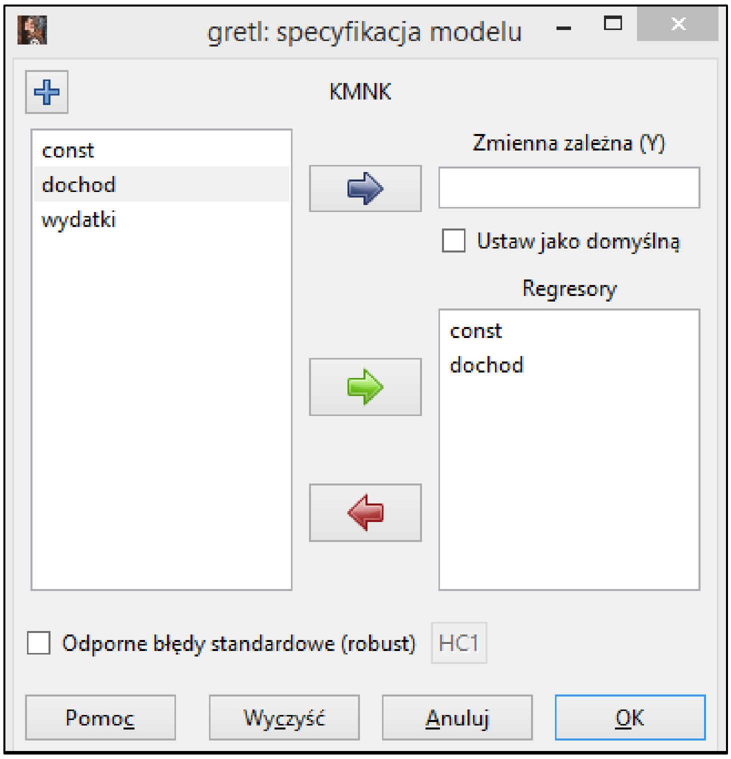
Oszacowania parametrów można dokonać ręcznie wyliczając wartości ze wzorów:

$$\hat{\alpha}_o = \bar{y} - \hat{\alpha}_1 * \bar{x}$$

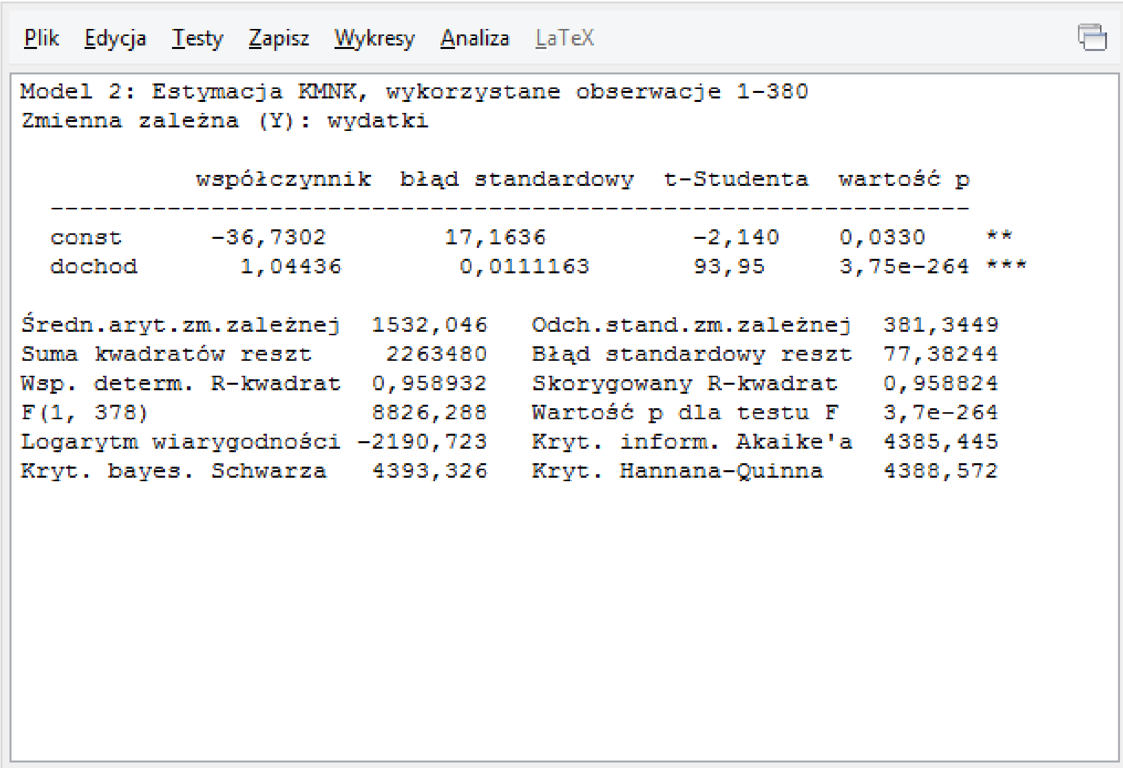
$$\hat{\alpha}_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$$

Z doświadczenia wiem, że czasochłonne jest ręczne liczenie tych wartości, dlatego pokażę Wam, jak te wartości wyliczyć w programie GRETL.

Zaznaczamy dane i klikamy Model -> Klasyczna metoda najmniejszych kwadratów. Pojawia się okno:



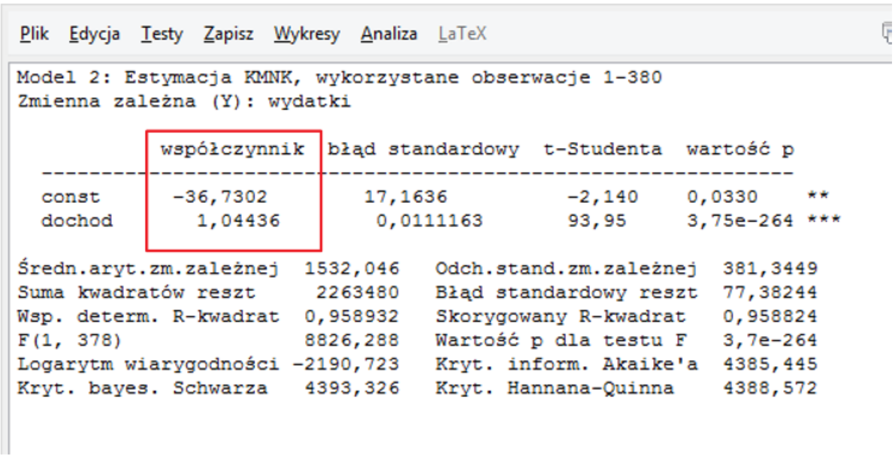
Ustawiamy zmienną zależną i regresory (zmienne niezależne). Const zostaje. Pojawi się okno modelu:




```
Model 2: Estymacja KMNK, wykorzystane obserwacje 1-380
Zmienna zależna (Y): wydatki

             współczynnik  błąd standardowy  t-Studenta  wartość p
  ---------------------------------------------------------------
  const      -36,7302       17,1636           -2,140     0,0330    **
  dochod       1,04436       0,0111163        93,95     3,75e-264 ***

Średn.aryt.zm.zależnej  1532,046   Odch.stand.zm.zależnej  381,3449
Suma kwadratów reszt     2263480   Błąd standardowy reszt  77,38244
Wsp. determ. R-kwadrat  0,958932   Skorygowany R-kwadrat   0,958824
F(1, 378)               8826,288   Wartość p dla testu F   3,7e-264
Logarytm wiarygodności -2190,723   Kryt. inform. Akaike'a  4385,445
Kryt. bayes. Schwarza   4393,326   Kryt. Hannana-Quinna    4388,572
```

Model ekonometryczny wygląda następująco:

$$wydatki = \alpha_0 + \alpha_1 * dochód + \varepsilon$$




```
Model 2: Estymacja KMNK, wykorzystane obserwacje 1-380
Zmienna zależna (Y): wydatki

             współczynnik  błąd standardowy  t-Studenta  wartość p
  ---------------------------------------------------------------
  const      -36,7302       17,1636           -2,140     0,0330    **
  dochod       1,04436       0,0111163        93,95     3,75e-264 ***

Średn.aryt.zm.zależnej  1532,046   Odch.stand.zm.zależnej  381,3449
Suma kwadratów reszt     2263480   Błąd standardowy reszt  77,38244
Wsp. determ. R-kwadrat  0,958932   Skorygowany R-kwadrat   0,958824
F(1, 378)               8826,288   Wartość p dla testu F   3,7e-264
Logarytm wiarygodności -2190,723   Kryt. inform. Akaike'a  4385,445
Kryt. bayes. Schwarza   4393,326   Kryt. Hannana-Quinna    4388,572
```

Zaznaczone dane to nasze oszacowane parametry. Zanim zapiszemy oszacowany model powinnyśmy najpierw ocenić istotność parametrów (o czym w następnym rozdziale). Jeżeli parametry będą istotne to nasz model będzie wyglądał następująco:

$$wydatki = -36{,}7302 + 1{,}044 * dochód$$

5. Ocena istotności parametrów

Wszystkie parametry muszą być w modelu istotne, co oznacza tyle że zmienne przy tych parametrach wpływają na naszą zmienną objaśnianą (Y). Sprawdzamy to testem istotności parametrów. Wyróżniamy test F i test t na istotność parametrów. Najpierw test t:

$$H_0: \alpha_i = 0 \qquad (parametr\ nieistotny)$$

$$H_1: \alpha_i \neq 0 \qquad (parametr\ istotny)$$

Statystyka testowa testu:

$$T = \frac{\hat{\alpha}_i}{S(\hat{\alpha}_i)}$$

$S(\hat{\alpha}_i) - błąd\ standardowy\ parametru$

$$H_0\ odrzucamy\ gdy\ T > t_{\frac{\alpha}{2}, n-k-1}$$

lub

$$H_0\ odrzucamy\ gdy\ p < \alpha$$

Test ten musimy przeprowadzić dla każdego parametru. Jeżeli choćby jeden jest nieistotny należy go wyeliminować z modelu i ponownie oszacować model.

Różnica między testem t a testem F jest taka, że test t pozwala ocenić istotność pojedynczych parametrów, natomiast test F weryfikuje wszystkie parametry naraz:

$$H_0: \alpha_1 = \alpha_2 = \cdots = \alpha_i = 0$$

$$H_1: \alpha_1 \neq \alpha_2 \neq \cdots \neq \alpha_i \neq 0$$

Statystyka testowa:

$$F = \frac{n-k-1}{k} * \frac{R^2}{1-R^2}$$

Gdzie:

k- liczba zmiennych objaśniających

R^2- współczynnik determinacji R^2

$$H_0\ odrzucamy\ gdy\ F > F_{k, n-k-1}$$

lub

$$H_0\ odrzucamy\ gdy\ p < \alpha$$

W pierwszym przypadku korzystamy z tablic Fishera Snedecora, w drugim z programu GRETL.

Przykład:

```
Plik  Edycja  Testy  Zapisz  Wykresy  Analiza  LaTeX

Model 2: Estymacja KMNK, wykorzystane obserwacje 1-380
Zmienna zależna (Y): wydatki

             współczynnik  błąd standardowy  t-Studenta  wartość p
  ---------------------------------------------------------------
  const       -36,7302        17,1636          -2,140     0,0330    **
  dochod        1,04436        0,0111163       93,95      3,75e-264 ***

Średn.aryt.zm.zależnej  1532,046   Odch.stand.zm.zależnej  381,3449
Suma kwadratów reszt     2263480   Błąd standardowy reszt  77,38244
Wsp. determ. R-kwadrat  0,958932   Skorygowany R-kwadrat   0,958824
F(1, 378)               8826,288   Wartość p dla testu F   3,7e-264
Logarytm wiarygodności -2190,723   Kryt. inform. Akaike'a  4385,445
Kryt. bayes. Schwarza   4393,326   Kryt. Hannana-Quinna    4388,572
```

$$wydatki = \alpha_0 + \alpha_1 * dochód + \varepsilon$$

Zbadaj istotność parametrów tego modelu.

Rozwiązanie

1) Hipotezy:

$$H_0: \alpha_0 = 0 \qquad (parametr\ nieistotny)$$

$$H_1: \alpha_0 \neq 0 \qquad (parametr\ istotny)$$

2) Statystyka testowa:

$$T = \frac{\hat{\alpha}_0}{S(\hat{\alpha}_0)} = \frac{-36,7302}{17,1636} = -2,1401$$

$$H_0\ odrzucamy\ gdy\ |T| > t_{\frac{\alpha}{2}, n-k-1}$$

$$t_{\frac{\alpha}{2}, n-k-1} = t_{\frac{0,05}{2}, 380-1-1} = t_{0,025, 378} = 1,966$$

$$2,1401 > 1,966$$

3) Wniosek:

Odrzucamy H_0. Parametr α_0 jest istotny statystycznie.

Istotność parametru α_1 zbadamy drugą wersją testu t-studenta.

1) Hipotezy:

$$H_0: \alpha_1 = 0 \qquad (parametr\ nieistotny)$$

$$H_1: \alpha_1 \neq 0 \qquad (parametr\ istotny)$$

2) H_0 $odrzucamy\ gdy\ p < \alpha$

Za wartość p przyjmujemy wartość znajdującą się w wierszu danej zmiennej (dochód) i kolumnie wartość p.

$$p = 3{,}75 * 10^{-264} < 0{,}05$$

3) Wniosek:

Odrzucamy H_0. Parametr α_0 jest istotny statystycznie.

Aby dopełnić obraz istotności parametrów, zbadam także ich istotność wykorzystując test F.

1) Hipotezy:

$$H_0: \alpha_0 = \alpha_1 = 0$$

$$H_1: \alpha_0 \neq \alpha_1 \neq 0$$

2) Statystyka testowa:

$$F = \frac{n-k-1}{k} * \frac{R^2}{1-R^2} = \frac{380-1-1}{1} * \frac{0{,}9589^2}{1-0{,}9589^2} = 4317{,}023$$

$$H_0\ odrzucamy\ gdy\ F > F_{k,n-k-1}$$

$$F_{1,380-1-1} = F_{1,378} = 3{,}866$$

$$4317{,}023 > 3{,}866$$

3) Wniosek

Odrzucamy H_0. Parametry w tym modelu są istotne statystycznie.

Jeżeli parametry w modelu są nieistotne (lub choćby jeden z nich) musimy je/go wyrzucić z modelu i ponownie zastosować klasyczną metodę najmniejszych kwadratów. Robimy to do momentu gdy wszystkie parametry są istotne. Gdy już to się stanie, możemy zapisać model oszacowany, z daszkami:

$$wydatki = -36{,}7302 + 1{,}044 * dochód$$

6. Interpretacja parametrów i zgodność z teorią ekonomii

Kolejnym etapem jest interpretacja parametrów. Tak jak już wspominałam mam dwa rodzaje parametrów: $\alpha_0\ i\ \alpha_1$. Są to parametry w równaniu z jedną zmienną objaśniającą. α_0 wskazuje na wartość zmiennej objaśnianej przy wartości 0 zmiennej objaśniającej. α_1 to współczynnik kierunkowy, który wskazuje na postać malejącą i rosnącą prostej tego równania. W matematyce finansowej jest coś takiego jak elastyczność zmiennej. Mówi ona o ile zmieni się Y jeżeli X zwiększymy o jedną jednostkę. Wartość elastyczności wykazuje pochodna Y.

$$Y = \alpha_0 + \alpha_1 * X$$

$Y' = \alpha_1$ (wynika to ze wzorów na pochodne)

$wydatki = -36{,}7302 + 1{,}044 * dochód$

Interpretacja α_0: Jeżeli dochody budżetowe na 1 mieszkańca będą na poziomie 0 zł to wydatki na 1 mieszkańca będą na poziomie -36,73 zł.

Interpretacja α_1: Jeżeli dochody budżetowe na 1 mieszkańca wzrosną o 1 zł to wydatki na 1 mieszkańca wzrosną o 1,04 zł.

Podstawową zasadą w budowie modelu ekonometrycznego jest jego zgodność z teorią ekonomii. Po co budować i analizować model, który do niczego nam się nie przyda. Najprościej zgodność tą wykazać poprzez właśnie interpretację parametrów. Najważniejszy jest współczynnik α_1. W naszym przypadku wzrost dochodów na 1 mieszkańca powoduje wzrost wydatków na 1 mieszkańca, co jest zgodne z teorią ekonomii.

Jeżeli mamy do czynienia z modelem regresji wielorakiej wówczas musimy zinterpretować także pozostałe współczynniki kierunkowe. Przykładowa interpretacja α_2::

Jeżeli dochody budżetowe na 1 mieszkańca wzrosną o 1 zł to wydatki na 1 mieszkańca wzrosną o 1,04 zł przy niezmienionych pozostałych czynnikach.

Musimy dodać „przy niezmienionych pozostałych czynnikach”, ponieważ jeżeli zmieni się inny parametr to nasze wydatki nie wzrosną o 1,04 zł, a np. o 2 zł.

7. Ocena dopasowania modelu

Model może być lepiej lub gorzej dopasowany do danych empirycznych. Są dwa parametry, które umożliwiają ocenę dopasowania modelu. Pierwszym z nich jest współczynnik determinacji R^2. Przyjmuje on wartości w zakresie $< 0{,}1 >$. Podobnie jak dla współczynnika korelacji, im wyższa jego wartość tym lepsze dopasowanie do danych.

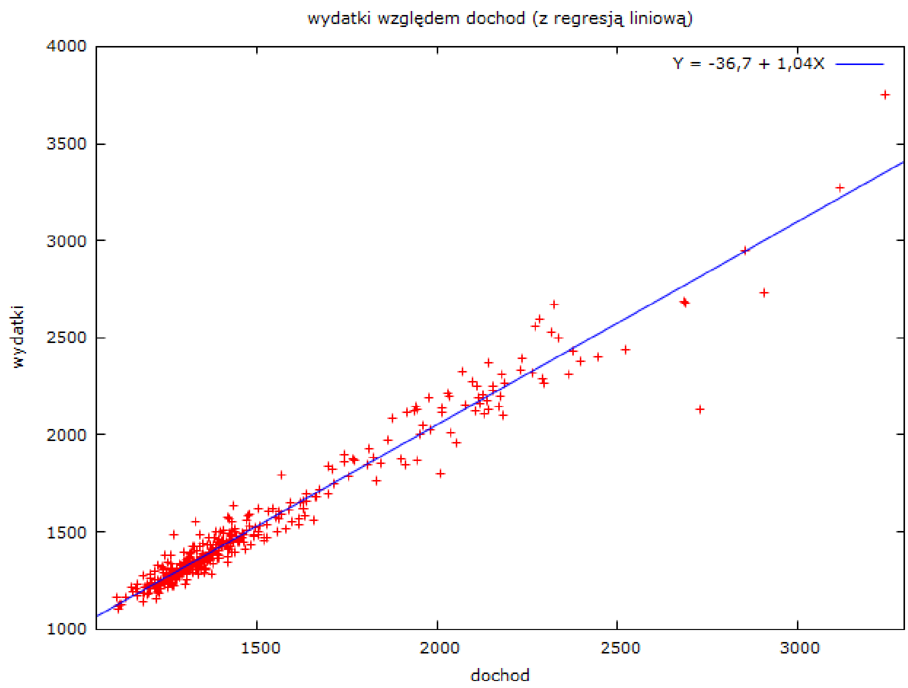
Przykład:

```
Model 2: Estymacja KMNK, wykorzystane obserwacje 1-380
Zmienna zależna (Y): wydatki

            współczynnik  błąd standardowy  t-Studenta  wartość p
  ---------------------------------------------------------------
  const      -36,7302        17,1636          -2,140     0,0330     **
  dochod       1,04436        0,0111163       93,95      3,75e-264 ***

Średn.aryt.zm.zależnej  1532,046  Odch.stand.zm.zależnej  381,3449
Suma kwadratów reszt     2263480  Błąd standardowy reszt  77,38244
Wsp. determ. R-kwadrat  0,958932  Skorygowany R-kwadrat   0,958824
F(1, 378)               8826,288  Wartość p dla testu F   3,7e-264
Logarytm wiarygodności -2190,723  Kryt. inform. Akaike'a  4385,445
Kryt. bayes. Schwarza   4393,326  Kryt. Hannana-Quinna    4388,572
```

Współczynnik determinacji wynosi 0,958932, co oznacza że 95,89% zmienności zmiennej wydatki budżetowe na 1 mieszkańca jest wyjaśniona poprzez liniowy związek ze zmienną dochody budżetowe na 1 mieszkańca.

Drugim parametrem jest błąd standardowy estymacji zwany także odchyleniem standardowym. Im błąd standardowy jest mniejszy tym lepiej przewidywany jest parametr, a tym samym cały model. Jeżeli zmienne posiadają duże błędy standardowe w porównaniu do wartości współczynnika tym oszacowanie prawdziwej, realnej wartości będzie mniej dokładne.

Ocena dopasowania modelu jest możliwa także poprzez wykresu dopasowania. W GRETLU znajdziemy go w zakładce Widok -> Wykresy zmiennych -> Wykres rozrzutu X-Y. W miejsce na osi X wpisujemy zmienną niezależną (dochód), natomiast w miejsce na osi Y wpisujemy zmienną zależną (wydatki). Otrzymujemy wykres rozrzutu:



W prawym górnym rogu otrzymaliśmy prostą najmniejszych kwadratów, nasz model ekonometryczny. Interpretacja tego wykresu może być następująca: Punkty układają się wzdłuż prostej, co świadczy o bardzo dobrym dopasowaniu modelu do danych empirycznych. Można posłużyć się także współczynnikiem korelacji Pearsona. Zaznaczamy dane, klikamy prawym przyciskiem myszy i wybieramy Macierz korelacji. Pierwsza wartość to współczynnik korelacji.

```
corr(dochod, wydatki) = 0,97925083
Hipoteza zerowa: R = 0, brak korelacji:
 t(378) = 93,9483, przy dwustronnym obszarze krytycznym p = 0,0000
```

8. Specyfikacja modelu i modele nieliniowe

W teorii ekonomii często zdarzało się, że zjawiska ekonomiczne nie układały się wzdłuż prostej, były to modele nieliniowe. Wyróżniamy 3 rodzaje modeli:

- Liniowe,
- Nieliniowe sprowadzalne do liniowych,
- Nieliniowe niesprowadzalne do liniowych.

MNK bada tylko modele liniowe, dlatego niezbędne jest określenie czy ta postać jest odpowiednia dla tego zjawiska. Służy do tego test na specyfikację. Specyfikacja umożliwia określenie czy dane zmienne w modelu są ok, czy mamy dobrą postać analityczną i określenie charakteru zależności modelu. Wykorzystujemy test Ramsey'a, inaczej zwany test RESET.

1) Hipotezy:

$$H_0: specyfikacja\ jest\ poprawna$$

$$H_1: specyfikacja\ nie\ jest\ poprawna$$

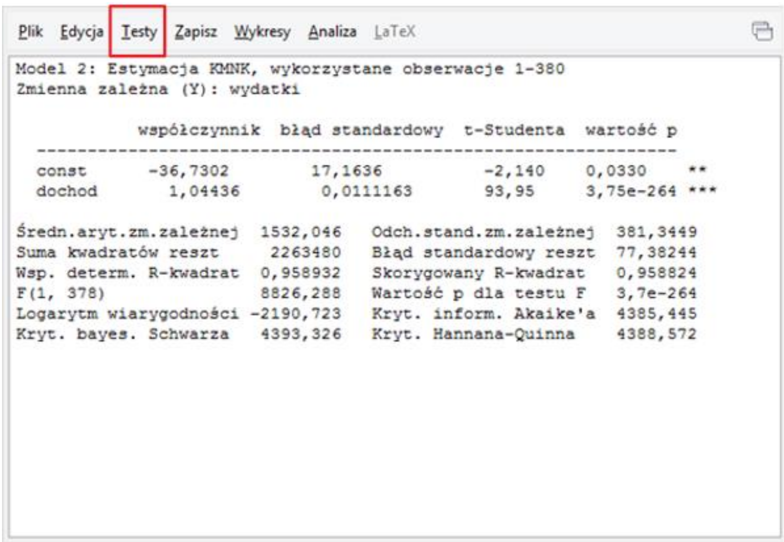
2) Statystyka testowa:

$$F = \frac{(R_p^2 - R^2)/2}{(1 - R_p^2)/(n-k)}$$

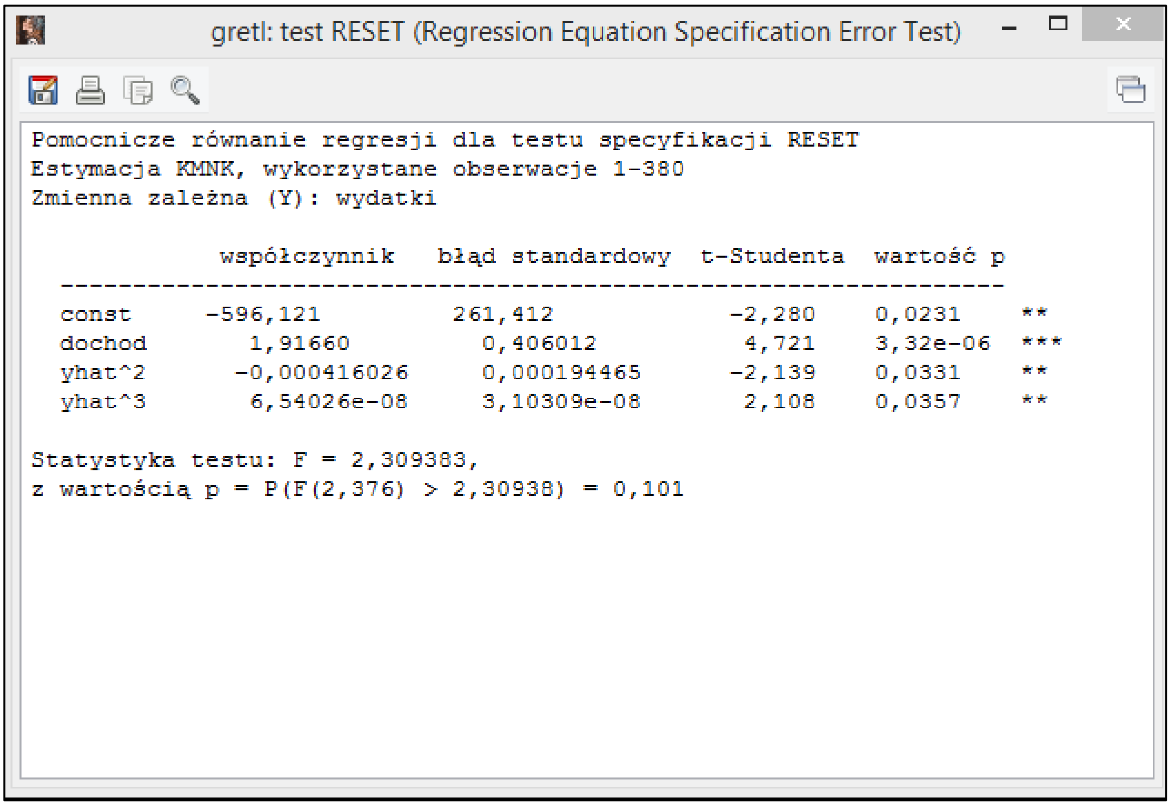
W teście tym wykorzystujemy tzw. regresję pomocniczą. Regresorami są wszystkie zmienne objaśniające oraz kwadrat i sześcian $\hat{y}$. Oznacza to tyle, że w przypadku liczby k będzie ona wynosiła tyle co liczba X + 2.

$$H_0\ odrzucamy\ gdy\ F > F_{2,n-k-1}$$

Również w przypadku tego testu nie musimy liczyć statystyki testowej ręcznie, ponieważ GRETL umożliwia jej wyliczenie. W ikonie modelu znajduje się zakładka Testy.



Wybieramy test specyfikacji RESET, następnie kwadrat i sześcian zmiennej i otrzymujemy wyniki:



gretl: test RESET (Regression Equation Specification Error Test)

```
Pomocnicze równanie regresji dla testu specyfikacji RESET
Estymacja KMNK, wykorzystane obserwacje 1-380
Zmienna zależna (Y): wydatki

             współczynnik   błąd standardowy  t-Studenta  wartość p
  ---------------------------------------------------------------
  const     -596,121         261,412            -2,280    0,0231    **
  dochod       1,91660         0,406012          4,721    3,32e-06  ***
  yhat^2      -0,000416026     0,000194465      -2,139    0,0331    **
  yhat^3       6,54026e-08     3,10309e-08       2,108    0,0357    **

Statystyka testu: F = 2,309383,
z wartością p = P(F(2,376) > 2,30938) = 0,101
```

Statystyka testu F wynosi 2,3094.

$$H_0 \; odrzucamy \; gdy \; F > F_{2,n-k-1}$$

$n = 380$

$k = 1 + 2 = 3$ $(jedna \; zmienna \; objaśniająca + zawsze \; 2)$

$F_{2,n-k-1} = F_{2,380-3-1} = F_{2,376} = 3{,}01973$

$$2{,}3094 < 3{,}01973$$

Nie odrzucamy H_0. Specyfikacja jest poprawna.

A co się dzieje, gdy specyfikacja jest niepoprawna? W pierwszej kolejności powinniśmy sprawdzić istotność korelacji (roz. 3). W drugiej kolejności istotność parametrów (roz. 5), a na sam koniec zastosować transformacje. Jak już pisałam wcześniej wyróżniamy modele nieliniowe sprowadzalne do liniowych. Aby sprowadzić je do modelu liniowego należy zastosować odpowiednią transformację. Poniższa tabela ukazuje wzór modelu nieliniowego i jego transformatę.

Funkcja	**Model nieliniowy**	**Transformata**
Wykładnicza	$Y = \alpha_0 + \alpha_1^X$	$\log Y = \tilde{\alpha}_0 + \tilde{\alpha}_1 * X$
Potęgowa	$Y = \alpha_0 * X^{\alpha_1}$	$\log Y = \tilde{\alpha}_0 + \tilde{\alpha}_1 * \log X$
Logarytmiczna	$Y = \alpha_0 + \alpha_1 * \log X$	$Y = \alpha_0 + \alpha_1 * \log X$
Hiperboliczna	$Y = \alpha_0 + \frac{\alpha_1}{X}$	$Y = \alpha_0 + \frac{\alpha_1}{X}$

9. Weryfikacja założeń MNK

W pierwszej kolejności wyjaśnię Wam, co to jest MNK. Jest to metoda, która polega na oszacowaniu parametrów modelu ekonometrycznego. Posługuje się ona wyznaczeniem minimum sumy kwadratów reszt. Metoda najmniejszych kwadratów wymaga jednak spełnienia 3 warunków:

- Równanie musi być liniowe lub nieliniowe sprowadzalne do liniowej postaci
- W próbie nie może być mniej obserwacji niż parametrów
- Zmienne są liniowo niezależne.

Dwa pierwsze warunki zostały już omówione. Co do trzeciego warunku to co to w ogóle znaczy? Jeżeli zmienne są zależne liniowo to jest to współliniowość zmiennych, czyli z wartości jednej zmiennej możemy wyznaczyć wartości pozostałych. Jest to bardzo niekorzystne zjawisko, dlatego niezbędne jest jego wyeliminowanie. Współliniowość można wykryć, jeżeli wszystkie parametry danego modelu są nieistotne, tzn. widoczne są duże błędy standardowe. Niezbędne jest wtedy sprawdzenie wskaźnika współliniowości VIF. Jeżeli jego wartość VIF dla danego parametru przewyższa 10 to ten parametr może być przyczyną współliniowości. W GRETLU wskaźnik VIF znajduje się także w zakładce Analiza, a następnie Ocena współliniowości.

```
Plik  Edycja  Testy  Zapisz  Wykresy  Analiza  LaTeX

Model 1: Estymacja KMNK, wykorzystane obserwacje 2002:1-2005:2 (N = 14)
Zmienna zależna (Y): PKB

                   współczynnik   błąd standardowy  t-Studenta  wartość p
  ----------------------------------------------------------------------
  const            -35391,2         26631,0          -1,329      0,2205
  spozycie              0,382513        0,326994      1,170      0,2758
  naklady_brutto        0,170556        0,283740      0,6011     0,5644
  popyt_krajowy         0,812793        0,472152      1,721      0,1235
  wartosc_dod_prz       0,0586803       0,632207      0,09282    0,9283
  wartosc_dod_us        0,109642        0,289851      0,3783     0,7151

Średn.aryt.zm.zależnej  208245,9   Odch.stand.zm.zależnej  18658,52
Suma kwadratów reszt    27018537   Błąd standardowy reszt  1837,748
Wsp. determ. R-kwadrat  0,994030   Skorygowany R-kwadrat   0,990299
F(5, 8)                 266,4132   Wartość p dla testu F   1,14e-08
Logarytm wiarygodności -121,1760   Kryt. inform. Akaike'a  254,3519
Kryt. bayes. Schwarza   258,1863   Kryt. Hannana-Quinna    253,9970
Autokorel.reszt - rho1  0,163171   Stat. Durbina-Watsona   1,563064
```

```
Ocena współliniowości VIF(j) - czynnik rozdęcia wariancji
VIF (Variance Inflation Factors) - minimalna możliwa wartość = 1.0
Wartości > 10.0 mogą wskazywać na problem współliniowości - rozdęcia wariancji

       spozycie   25,265
 naklady_brutto   50,340
  popyt_krajowy  251,288
wartosc_dod_prz   57,805
 wartosc_dod_us   20,030

VIF(j) = 1/(1 - R(j)^2), gdzie R(j) jest współczynnikiem korelacji wielorakiej
pomiędzy zmienną 'j' a pozostałymi zmiennymi niezależnymi modelu.

Belsley-Kuh-Welsch collinearity diagnostics:

                          --- variance proportions ---
    lambda      cond      const  spozycie naklady_~ popyt_kr~ wartosc_~ wartosc_~
     5,919     1,000      0,000     0,000     0,000     0,000     0,000     0,000
     0,071     9,109      0,000     0,000     0,022     0,000     0,000     0,000
     0,007    28,385      0,015     0,000     0,003     0,000     0,024     0,001
     0,002    48,843      0,034     0,000     0,005     0,000     0,004     0,099
     0,000   222,248      0,430     0,713     0,133     0,000     0,238     0,131
     0,000   550,776      0,520     0,287     0,837     1,000     0,734     0,769

  lambda = eigenvalues of X'X, largest to smallest
  cond   = condition index
  note: variance proportions columns sum to 1.0
```

W tym przypadku zmienną, która może wywoływać współliniowość jest zmienna popyt_krajowy. Jak zapobiec temu zjawisku? Występuje ono najczęściej przy małej próbie. Aby zapobiec jej skutkom, należy zwiększyć próbę.

Założenia MNK badają składnik losowy. Jest to nasz ε w równaniu, zwany także zaburzeniem losowym, zakłóceniami, błędem losowym czy resztami. Przykładowo jeżeli badamy jak podaż wpływa na popyt to wiadomo, że na popyt wpływają także inne zmienne. Ten epsilon wyraża inne zmienne, które nie zostały ujęte w modelu. Wyróżniamy 5 głównych założeń MNK:

I. Zmienne niezależne są nieskorelowane z zakłóceniami losowymi
II. Średnia wartość zakłóceń jest równa 0
III. Błąd losowy jest homoscedastyczny
IV. Brak autokorelacji składnika losowego
V. Składnik losowy posiada rozkład normalny

I. Zmienne niezależne są nieskorelowane z zakłóceniami losowymi

Posługujemy się tym samym testem, co do istotności korelacji, jednak w hipotezach bierzemy pod uwagę zmienne objaśniające oraz reszty.

$$H_0: \rho(X, \varepsilon) = 0$$

$$H_1: \rho(X, \varepsilon) \neq 0$$

Statystyka testowa:

$$T = \frac{|r| * \sqrt{n-2}}{\sqrt{1-r^2}}$$

$$H_0 \ odrzucamy \ gdy \ T > t_{\frac{\alpha}{2}, n-2}$$

II. Średnia wartość zakłóceń jest równa 0

Założenie to jest spełnione gdy w modelu mamy wyraz wolny. Jeżeli go nie ma, nie możemy stwierdzić, że założenie nie jest spełnione. Musimy zrobić test.

$$H_0: \mu(\varepsilon) = 0$$

$$H_1: \mu(\varepsilon) \neq 0$$

Statystyka testowa:

$$Z = \frac{\bar{X} - \mu}{S} * \sqrt{n}$$

$$H_0 \ odrzucamy \ gdy \ |Z| > z_{\frac{\alpha}{2}}$$

lub

$$T = \frac{\bar{X} - \mu}{S} * \sqrt{n}$$

$$H_0 \ odrzucamy \ gdy \ |T| > t_{\frac{\alpha}{2}, n-1}$$

Jeżeli nasza próba przekracza 30 wówczas wybieramy pierwszą statystykę, co wiąże się z zastosowaniem tablic rozkładu normalnego. Jeżeli próba jest mniejsza niż 30, stosujemy drugą statystykę i rozkładu t-studenta.

III. Składnik losowy jest homoscedastyczny

Oznacza to że wariancja składnika losowego jest stała. Weryfikujemy następujące hipotezy:

$$H_0: \sigma^2 = const \ (homoskedastyczność)$$

$$H_1: \sigma^2 \neq const \ \ (heteroskedastyczność)$$

W przypadku tego założenia mamy do wyboru 3 testy:

- test Goldfielda-Quandta
- test Breusch'a-Pagana
- test White'a.

Test Goldfielda-Quandta stosujemy gdy wiemy, które zmienne wywołują heteroskedastyczność i jest ich więcej niż jedna. Test Breusch'a-Pagana stosujemy gdy wiemy, która zmienna wywołuje heteroskedastyczność i jest to dokładnie jedna zmienna. Test White'a jest na końcu, ponieważ jest on najsłabszym testem. Stosujemy go, gdy nie wiemy, która zmienna może wywoływać heteroskedastyczność. Po czym poznamy, która to zmienna? Po wykresie między resztami a np. czasem.

IV. Nie występuje autokorelacja składnika losowego

Założenie to dotyczy wyłącznie szeregów czasowych, czyli modeli, w który uwzględniany jest czas. Brak autokorelacji składnika losowego dotyczy korelacji między poszczególnymi zaburzeniami. Autokorelacja pierwszego rzędu bada zaburzenia sąsiadujące, natomiast autokorelacja wyższych rzędów- zaburzenia oddalone od siebie o wysokość rzędu.

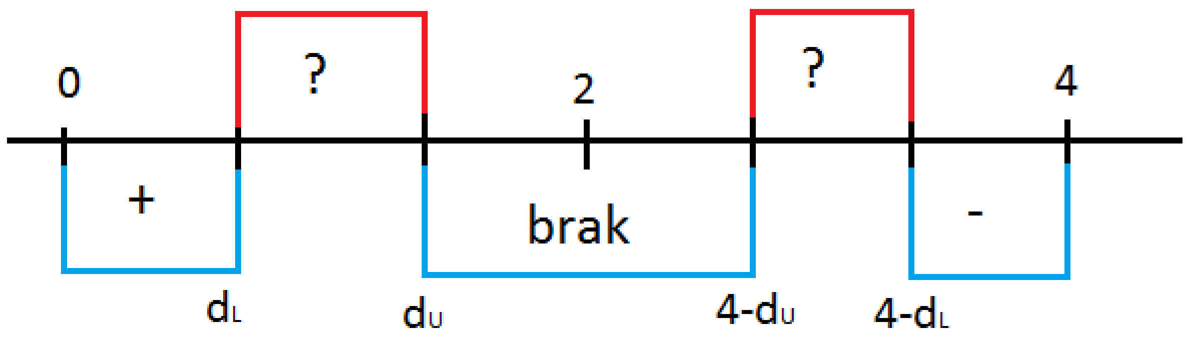
Pierwsze co robimy gdy patrzymy na model to sprawdzamy czy jest to szereg czasowy. Jeżeli nie, to przechodzimy do założenia nr 5. Jeżeli tak to musimy sprawdzić czy w modelu znajduje się wyraz wolny. Jest to warunek konieczny do testu na autokorelację I rzędu. Jeżeli jest, to robimy test DW. Jeżeli nie ma, to robimy test mnożników Lagrangre'a (test dla wyższych rzędów).

To od początku. Aby zastosować test Durbina-Watsona na autokorelację I rzędu muszą być spełnione dwa warunki:

- w modelu występuje wyraz wolny,
- w modelu nie występują zmienne opóźnione.

$$H_0: \rho(e_t, e_{t-1}) = 0 \qquad (brak\ autokorelacji)$$
$$H_1: \rho(e_t, e_{t-1}) \neq 0 \quad (istnieje\ autokorelacja)$$

Do weryfikacji tego testu potrzebujemy 3 danych: statystyki DW oraz dwóch liczb odczytanych z tablic Durbina-Watsona $d_L\ i\ d_U$. Wartość statystyki DW otrzymamy w zakładce Testy -> Test Durbina-Watsona. Następnie musimy odczytać wartości z tablic, a do tego potrzebujemy dwóch danych: n- liczba obserwacji oraz k- liczba zmiennych objaśniających bez wyrazu wolnego. Następnie rysujemy schemat.



+ oznacza autokorelację dodatnią

- oznacza autokorelację ujemną

brak oznacza brak autokorelacji (założenie jest spełnione)

? oznacza że występuje tzw. obszar niekonkluzywności, tzn. nie wiemy czy występuje autokorelacja

Nawet jeżeli wyjdzie nam autokorelacja w teście DW to i tak badamy autokorelację wyższych rzędów, o ile oczywiście mamy niezbędne statystyki. Nie badamy wszystkich rzędów. Badamy rząd, które odpowiada częstotliwości danych w roku i jego wielokrotność. Tzn. jeżeli dane są półroczne to najniższy rząd jaki badamy to 2, 4, 6 itd. Kwartalne- 4, 8, 12 itd. Miesięczne- 12, 24, 36 itd. Przeprowadzamy tyle testów, ile mamy danych.

Test mnożników Lagrangre'a opiera się na hipotezach:

$$H_0: \rho_1 = 0, \rho_2 = 0, \ldots, \rho_p = 0$$

$$H_1: \rho_1 \neq 0, \rho_2 \neq 0, \ldots, \rho_p \neq 0$$

Statystyka testowa:

$$LM = (n - p)R_{\hat{u}}^2$$

$$H_0 \; odrzucamy \; gdy \; LM > \chi_p^2$$

Do tego testu wykorzystujemy tablice chi-kwadrat, gdzie p jest rzędem badanej autokorelacji. W GRETLU wartość statystyki otrzymamy w zakładce Testy -> Testy autokorelacji. W kolejnym kroku musimy wybrać rząd badanej autokorelacji.

V. Składnik losowy ma rozkład normalny

Do badania rozkładu normalnego wykorzystujemy test Shapiro-Wilka. Hipotezy:

$$H_0: \varepsilon \sim N(0, \delta^2)$$

$$H_1: \varepsilon \nsim N(0, \delta^2)$$

Wartość SW otrzymamy w GRETLU w zakładce Zmienne -> Testy normalności rozkładu.

$$H_0 \; odrzucamy \; gdy \; SW < SW^*_{n,\alpha}$$

W tym teście wykorzystujemy tablice Shapiro-Wilka, gdzie n- liczba obserwacji, a $\alpha = 0{,}05$.

10. Zadanie

Zbadano liczbę nowo udzielonych kredytów mieszkaniowych (w tys.) w zależności od średniej ceny mieszkania za m2 (w tys. zł.), średniego miesięcznego wynagrodzenia (w tys. zł) oraz indeksu dostępności mieszkań M3.

Model 1: Estymacja KMNK, wykorzystane obserwacje 2007:1-2014:4 (N = 32)
Zmienna zależna (Y): kredyty

	Współczynnik	Błąd stand.	t-Studenta	wartość p
cena	-8,88247	0,843406	-10,5317	<0,00001
wynagrodzenia	1,4111	0,167007	8,4494	<0,00001
indeks M3	3,66427e-05	1,23943e-05	2,9564	0,00613

Średn.aryt.zm.zależnej	11,31253	Odch.stand.zm.zależnej	2,422761
Suma kwadratów reszt	127,8729	Błąd standardowy reszt	2,099860
Wsp. determ. R-kwadrat	0,970103	Skorygowany R-kwadrat	0,968041
F(3, 29)	313,6652	Wartość p dla testu F	3,42e-22
Logarytm wiarygodności	-67,57085	Kryt. inform. Akaike'a	141,1417
Kryt. bayes. Schwarza	145,5389	Kryt. Hannana-Quinna	142,5993
Autokorel.reszt - rho1	0,738152	Stat. Durbina-Watsona	0,525895

Otrzymano także pozostałe testy:

Test RESET na specyfikację -
Hipoteza zerowa: specyfikacja poprawna
Statystyka testu: F= 2,4468

Test White'a na heteroskedastyczność reszt
Statystyka testu: LM = 15,9959

Test Breuscha-Pagana na heteroskedastyczność
Statystyka testu: LM = 4,51247
z wartością p = P(Chi-kwadrat(3) > 4,51247) = 0,21118

Test na normalność rozkładu reszt -
Statystyka Shapiro-Wilka = 0,959707

cena	wynagrodzenia	indeks M3	reszty	
1,0000	-0,9453	-0,2236	-0,1920	cena
	1,0000	0,2104	0,0414	wynagrodzenia
		1,0000	0,1024	indeks M3
			1,0000	reszty


```
Statystyki opisowe
Zmienna: reszty

Średnia          0,11093
Odch.stand.      2,02786
Wsp.zmienności   18,2799
```

Polecenie:

Zweryfikuj założenia MNK.

Rozwiązanie

W pierwszej kolejności musimy zweryfikować postać modelu, istotność parametrów oraz specyfikację.

$$kredyty = \alpha_1 * cena + \alpha_2 * wynagrodzenia + \alpha_3 * indeks\ M3 + \varepsilon$$

Istotność parametrów

$$H_0: \alpha_i = 0 \qquad (parametr\ nieistotny)$$

$$H_1: \alpha_i \neq 0 \qquad (parametr\ istotny)$$

$$H_0\ odrzucamy\ gdy\ p < \alpha$$

Dla α_1: wartość $p =< 0{,}00001 \ < 0{,}05$ Odrzucamy H_0. Parametr α_1 jest istotny.

Dla α_2: wartość $p =< 0{,}00001 \ < 0{,}05$ Odrzucamy H_0. Parametr α_2 jest istotny.

Dla α_3: wartość $p = 0{,}00613 \ < 0{,}05$ Odrzucamy H_0. Parametr α_3 jest istotny.

Wszystkie parametry są istotne, dlatego możemy zapisać oszacowaną postać modelu.

$$kredyty = -8{,}88 * cena + 1{,}41 * wynagrodzenia + 3{,}66 * 10^{-5} * indeks\ M3$$

Interpretacja parametrów

α_1: Jeżeli średnia wartość ceny mieszkania za m2 wzrośnie o 1 tys. zł to liczba nowo udzielonych kredytów mieszkaniowych spadnie o 8,88 tys.

α_2: Jeżeli średnie miesięczne wynagrodzenie wzrośnie o 1 tys. zł to liczba nowo udzielonych kredytów mieszkaniowych wzrośnie o 1 411.

α_3: Jeżeli indeks dostępności mieszkań M3 wzrośnie o 1 % to liczba nowo udzielonych kredytów mieszkaniowych wzrośnie o $3{,}66 * 10^{-5}$ tys.

Ocena dopasowania modelu

Współczynnik determinacji R^2 wynosi 0,970103, co oznacza, że 97,01% zmienności liczby nowo udzielonych kredytów mieszkaniowych jest wyjaśnione poprzez liniowy związek ze zmiennymi objaśniającymi, czyli średnią wartością ceny mieszkania za m2, średnim miesięcznym wynagrodzeniem oraz indeksem dostępności mieszkań M3.

Ocena specyfikacji modelu

$H_0: specyfikacja\ jest\ poprawna$

$H_1: specyfikacja\ nie\ jest\ poprawna$

$F = 2,4468$

$$H_0\ odrzucamy\ gdy\ F > F_{2,n-k-1}$$

$F_{2,n-k-1} = F_{2,32-5-1} = F_{2,26} = 3,36$

$2,4468 < 3,36$

Nie odrzucamy H_0. Specyfikacja jest poprawna.

Weryfikacja założeń MNK

I. *Zmienne niezależne są nieskorelowane z zakłóceniami losowymi*

$$H_0: \rho(X, \varepsilon) = 0$$

$$H_1: \rho(X, \varepsilon) \neq 0$$

Statystyka testowa:

$$T = \frac{|r| * \sqrt{n-2}}{\sqrt{1-r^2}}$$

$$H_0\ odrzucamy\ gdy\ T > t_{\frac{\alpha}{2},n-2}$$

Dla α_1: $r = -0,1920$ $T = 1,071564$

Dla α_2: $r = 0,0414$ $T = 0,226952$

Dla α_3: $r = 0,1024$ $T = 0,563832$

$t_{\frac{\alpha}{2},n-2} = t_{0,025,30} = 2,04$

Dla wszystkich zmiennych objaśniających warunek $T > t_{\frac{\alpha}{2},n-2}$ nie został spełniony. Nie odrzucamy H_0. Zmienne niezależne są nieskorelowane z zakłóceniami losowymi. Założenie MNK zostało spełnione.

II. *Średnia wartość zakłóceń jest równa 0*

W modelu nie mamy wyrazu wolnego, w związku z czym niezbędne jest przeprowadzenie testu wartości zakłóceń. Liczba obserwacji wynosi 32, dlatego stosujemy pierwszą wartość statystyki i korzystamy z tablic rozkładu normalnego.

$$H_0: \mu(\varepsilon) = 0$$
$$H_1: \mu(\varepsilon) \neq 0$$

Statystyka testowa:

$$Z = \frac{\bar{X} - \mu}{S} * \sqrt{n} = \frac{0,11093 - 0}{2,02786} * \sqrt{32} = 0,309447$$

$$H_0 \; odrzucamy \; gdy \; |Z| > \; z_{\frac{\alpha}{2}}$$

$z_{\frac{\alpha}{2}} = 1,96$ $\qquad\qquad$ $0,309447 < 1,96$

Nie odrzucamy H_0. Średnia wartość zakłóceń jest równa 0. Założenie MNK zostało spełnione.

III. Błąd losowy jest homoscedastyczny

Mamy wykres zależności między resztami a czasem i widać na nim zależność między tymi zmiennymi. Oznacza to że wiem jaka zmienna może wywoływać heteroskedastyczność i jest dokładnie jedna. Stosujemy test Breusch'a-Pagana.

$$H_0: \sigma^2 = const \; (homoskedastyczność)$$
$$H_1: \sigma^2 \neq const \; (heteroskedastyczność)$$
$$H_0 \; odrzucamy \; gdy \; p < 0,05$$
$$p = 0,21118 > 0,05$$

Odrzucamy H_0. Występuje heteroskedastyczność. Założenie nie jest spełnione.

IV. Brak autokorelacji składnika losowego

Jest to szereg czasowy, ale nie posiada wyrazu wolnego. Oznacza to tyle, że badamy to założenie, jednak nie możemy zrobić testu Durbina-Watsona (mimo podanej statystyki DW). Nie posiadamy także testów na autokorelację wyższych rzędów. Z tych przyczyn nie możemy określić czy to założenie jest spełnione.

V. Składnik losowy posiada rozkład normalny

$$H_0\colon \varepsilon \sim N(0, \delta^2)$$
$$H_1\colon \varepsilon \nsim N(0, \delta^2)$$
$$H_0 \; odrzucamy \; gdy \; SW < SW_{n,\alpha}^*$$

$SW = 0{,}959707$

$SW_{n,\alpha}^* = SW_{32,0{,}05}^* = 0{,}927$

$$0{,}959707 > 0{,}927$$

Nie odrzucamy H_0. Składnik losowy posiada rozkład normalny. Założenie zostało spełnione.