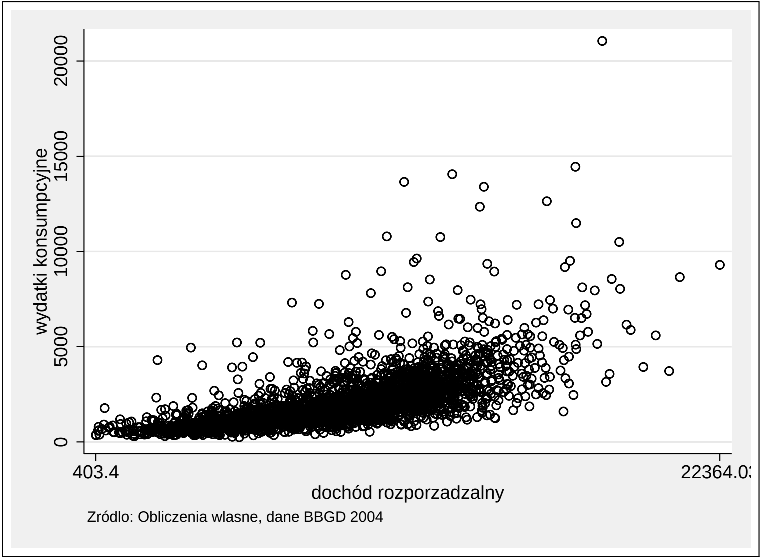
Heteroscedastyczność


wydatki konsumpcyjne
0
5000
10000
15000
20000
403.4
22364.0
dochód rozporzadzalny
Zródlo: Obliczenia wlasne, dane BBGD 2004

- Model zjawiska

$$y = X\beta + \varepsilon$$

- Wariancja składnika losowego

$$var(\varepsilon) = \Omega = \sigma^2\mathbb{V}$$

- W przypadku heteroscedastycznego składnika losowego i braku autokorelacji $\mathbb{V}$ jest macierzą diagonalną

1. Wariancja może być funkcją liniową zmiennych macierzy Z

$$var(\varepsilon_i) = \sigma_i^2 = \delta z_i \quad \delta > 0$$

2. Forma kwadratowa zabezpiecza przed ujemnością wariancji

$$var(\varepsilon_i) = \sigma_i^2 = \delta z_i^2 \quad \delta > 0$$

3. Wariancja może być również funkcją afiniczną zmiennych macierzy Z

$$var(\varepsilon_i) = \sigma_i^2 = \delta_1 + \delta_2 z_i^2 \quad \delta_1 > 0, \delta_2 > 0$$

W takim przypadku mówimy o heteroscedastyczności addytywnej.

4. Wariancja może przyjmować postać wykładniczą

$$var(\varepsilon_i) = \sigma_i^2 = exp(\delta_1 + \delta_2 z_i^2)$$

Wtedy mówimy o wariancji z heteroscedastycznością multiplikatywną.

5. W modelu może również występować wariancja przełącznikowa (*switching*)

$$var(\varepsilon_i) = \sigma_i^2 = \begin{cases} \sigma_1^2 & \text{dla i=1...s} \\ \sigma_2^2 & \text{dla i=s+1...T} \end{cases}$$

Ten typ wariancji może być połączony z każdą z uprzednio przedstawionych postaci.

Nieobciążoność

- Nieobciążoność estymatora wektora parametrów

$$E(b) = E(X'X)^{-1}X'y = E(X'X)^{-1}X'(X\beta + \varepsilon) =$$

$$= \underbrace{(X'X)^{-1}X'X}_{\mathbb{I}} E(\beta) + (X'X)^{-1}X' \underbrace{E(\varepsilon)}_{0} = \beta$$

- Estymator MNK jest nieobciążony, gdyż jego postać nie zależy od składnika losowego

Brak zgodności

- Brak zgodności estymatora wariancji wektora parametrów

$$var(b) = E(b-E(b))(b-E(b))' = E[(X'X)^{-1}X'\varepsilon\varepsilon'X(X'X)^{-1}]$$

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

$$var(b) = \sigma^2(X'X)^{-1}X'\mathbb{V}X(X'X)^{-1}$$

- Standardowe testy oparte na statystykach t, χ^2 oraz F będą wskazywać nieprawidłowe wyniki

Twierdzenie o faktoryzacji macierzy

Każdą dodatnio określoną macierz A można przedstawić w postaci

$$A = C\Lambda C^{-1}$$

gdzie kolumny macierzy C zawierają wektory własne macierzy A, a Λ jest macierzą diagonalną z wartościami własnymi macierzy A na diagonali.

- Zatem można przedstawić macierz wariancji jako

$$\Omega = C\Lambda C^{-1}$$

- Niech $\Lambda = \Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}}$
- Niech $P' = C\Lambda^{-\frac{1}{2}}$. Wobec tego $\Omega^{-1} = P'P$, ponieważ P (jeśli nie ma autokorelacji) i Λ są diagonalne to $P = P'$ oraz $\Lambda = \Lambda'$.
- Mnożąc model regresji przez macierz P z lewej strony uzyskujemy

$$Py = PX\beta + P\varepsilon$$

lub alternatywnie

$$y_\star = X_\star\beta + \varepsilon_\star \tag{1}$$

- Wariancja składnika losowego dla przekształconego modelu wynosi

$$E[\varepsilon_\star\varepsilon_\star'] = P\varepsilon\varepsilon'P' = \sigma^2 P\mathbb{V}P' = \sigma^2\mathbb{I}$$

Estymator UMNK wektora parametrów

- Zapiszmy wzór na estymator MNK

$$\hat{\beta} = (X_\star' X_\star)^{-1} X_\star *' y_\star$$

- Podstawiając otrzymujemy

$$\hat{\beta} = (X'P'PX)^{-1} X'P'Py$$

- W rezultacie

$$\hat{\beta} = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}y$$

Estymator UMNK wariancji wektora parametrów

- W MNK $var(b) = \sigma^2(X'X)^{-1}$
- W UMNK $var(\hat{\beta}) = \sigma^2(X'\mathbb{V}^{-1}X)^{-1}$
- Zatem wariancja estymatora UMNK nie jest równa wariancji estymatora MNK

$$var(b) = \sigma^2(X'X)^{-1}X'\mathbb{V}X(X'X)^{-1} \neq \sigma^2(X'\mathbb{V}^{-1}X)^{-1}$$

- Więc wykorzystując MNK nie można w prawidłowy sposób oszacować wariancji jeśli składnik losowy nie jest sferyczny

Twierdzenie Aitkena

Estymator Uogólnionej Metody Najmniejszych Kwadratów jest najlepszym, liniowym i nieobciążonym estymatorem wektora parametrów β o ile postać macierzy wariancji jest znana

- Twierdzenie Aitkena wynika wprost z zastosowania Twierdzenia Gaussa-Markowa do przekształconego modelu

- Praktyka. Nie znamy postaci macierzy $\mathbb{V}$
- Nie możemy jej oszacować ze względu na duży wymiar
- Szczególny przypadek - Ważona Metoda Najmniejszych Kwadratów

- Estymator UMNK dany jest wzorem:

$$\hat{\beta} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y$$

- Jeżeli nie ma autokorelacji, elementy ω_i (wagi) są znane to Ω^{-1} jest macierzą diagonalną o elementach na diagonali równych $1/\omega_i$.
- Przekształcając model mnożąc przez macierz P daną wzorem:

$$P = \begin{bmatrix} 1/\sqrt{\omega_1} & & 0 \\ & \ddots & \\ 0 & & 1/\sqrt{\omega_N} \end{bmatrix}$$

i stosując MNK do przekształconego modelu
- uzyskiwany jest estymator ważonej MNK

$$b = \left(\sum_{i=1}^{n} w_i x_i' x_i\right)^{-1}\left(\sum_{i=1}^{n} w_i x_i' y_i\right)$$

w którym $w_i = 1/\omega_i$.

- Zakłada się, że wariancja jest funkcją zmiennych Z

$$E(\varepsilon_i \mid Z_i) = \sigma_i^2 = \sigma^2 f(z_i)$$

- Wówczas odwrotność macierzy wariancji-kowariancji ma postać

$$\Omega^{-1} = \frac{1}{\sigma^2}\mathbb{V}^{-1} = \frac{1}{\sigma^2}\begin{bmatrix} \frac{1}{f(\alpha_0+\alpha z_i)} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{f(\alpha_0+\alpha z_i)} \end{bmatrix} = \sigma^2 L'L$$

SUMNK - algorytm

1. Szacowane są parametry modelu $y_i = x_i b + e_i$ i zapamiętywany wektor reszt.
2. Przeprowadzana jest regresja e_i^2 na stałą i wektor z_i
3. Szacowana jest macierz L.
4. Przekształcany jest za pomocą oszacowania $\hat{L}$ oryginalny model
5. Obliczana jest wartość estymatora

- W praktyce oszacowania uzyskane za pomocą SUMNK są zbliżone do oszacowań MNK.

SUMNK - przykład

```
  Source |       SS           df       MS      Number of obs   =    31,720
---------+----------------------------------   F(1, 31718)     =  20697.02
   Model |  3.8684e+10         1  3.8684e+10   Prob > F        =    0.0000
Residual |  5.9283e+10    31,718  1869075.24   R-squared       =    0.3949
---------+----------------------------------   Adj R-squared   =    0.3948
   Total |  9.7968e+10    31,719  3088609.94   Root MSE        =    1367.1

------------------------------------------------------------------------
   dochg |      Coef.   Std. Err.        t     P>|t|     [95% Conf. Interval]
---------+--------------------------------------------------------------
    kons |     0,7315     0,0051    143,86     0,00       0,7215     0,7414
   _cons |   789,7997    12,7227     62,08     0,00     764,8628   814,7366
------------------------------------------------------------------------
```

SUMNK - przykład

```
. predict e, resid
. gen e2=e^2
. reg e2 los, cformat(%9,4f) pformat(%5,2f) sformat(%8,2f)

   Source |       SS           df       MS      Number of obs   =    31,720
----------+----------------------------------   F(1, 31718)     =     16.62
    Model |  4.0371e+16         1  4.0371e+16   Prob > F        =    0.0000
 Residual |  7.7024e+19    31,718  2.4284e+15   R-squared       =    0.0005
----------+----------------------------------   Adj R-squared   =    0.0005
    Total |  7.7064e+19    31,719  2.4296e+15   Root MSE        =   4.9e+07

------------------------------------------------------------------------
       e2 |      Coef.   Std. Err.       t     P>|t|    [95% Conf. Interval]
----------+-------------------------------------------------------------
      los |   7,23e+05   1,77e+05     4,08     0,00    3,76e+05   1,07e+06
    _cons |  -3,58e+05   6,12e+05    -0,58     0,56   -1,56e+06   8,42e+05
------------------------------------------------------------------------
```

SUMNK - przykład

```
. gen w_os = kons/sqrt(fit)
. gen d_os = dochg/sqrt(fit)
. reg w_os d_os, cformat(%9,4f) pformat(%5,2f) sformat(%8,2f)

      Source |       SS           df       MS      Number of obs   =    31,720
-------------+----------------------------------   F(1, 31718)     =  19114.72
       Model |  17556.3773         1  17556.3773   Prob > F        =    0.0000
    Residual |  29132.1657    31,718  .918474232   R-squared       =    0.3760
-------------+----------------------------------   Adj R-squared   =    0.3760
       Total |   46688.543    31,719  1.47194246   Root MSE        =    .95837

------------------------------------------------------------------------------
        w_os |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        d_os |     0,5821     0,0042   138,26   0,00       0,5738      0,5903
       _cons |     0,5605     0,0093    60,39   0,00       0,5423      0,5787
------------------------------------------------------------------------------
```

Estymator odporny na heteroscedastyczność

- Jeżeli macierz wariacji-kowariancji Ω byłaby znana, wówczas estymatorem asymptotycznej macierzy kowariancji b byłoby

$$\Omega = \frac{1}{N}\left(\frac{1}{N}X'X\right)^{-1}\left(\frac{1}{N}X'\mathbb{V}^{-1}X\right)\left(\frac{1}{N}X'X\right)^{-1}$$

- Aby oszacować elementy środkowej macierzy trzeba znaleźć elementy macierzy

$$plim\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_{ij}^{2}x_i x_j'$$

- Ponieważ estymator MNK dla wektora parametrów β jest zgodny, estymatory dla reszt e_i są punktowo zgodne
- White (1980) pokazał, że

$$plim\frac{1}{N}\sum_{i=1}^{N}e_i^2 = x_i x_i' = plim\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_{ij}^{2}x_i x_j'$$

Estymator odporny na heteroscedastyczność

- Zatem następujące macierze są asymptotycznie równoważne

$$\frac{1}{N}X'\Omega X = \frac{1}{N}X'\sigma_i^2\mathbb{I}X$$

- Zgodnym estymatorem macierzy wariancji-kowariancji wektora parametrów jest

$$\hat{\Sigma}_b = n(X'X)^{-1}\left[\frac{1}{N}\sum e_i^2 x_i x_i'\right](X'X)^{-1}$$

Regresja odporna - przykład

```
. reg l_kons ldochg [aw=w04g], cformat(%9,4f) pformat(%5,2f) sformat(%8,2f)
(sum of wgt is 32,049.6321299076)

  Source |       SS           df       MS      Number of obs   =    32,054
---------+----------------------------------   F(1, 32052)     =  46832.70
   Model |   7183.9271         1   7183.9271   Prob > F        =    0.0000
Residual |  4916.63323    32,052  .153395521   R-squared       =    0.5937
---------+----------------------------------   Adj R-squared   =    0.5937
   Total |  12100.5603    32,053  .377517247   Root MSE        =    .39166

--------------------------------------------------------------------------
  l_kons |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
---------+----------------------------------------------------------------
  ldochg |     0,7296     0,0034   216,41    0,00      0,7230      0,7362
   _cons |     1,9129     0,0252    75,92    0,00      1,8635      1,9623
--------------------------------------------------------------------------
```

Regresja odporna - przykład

```
. reg l_kons ldochg [aw=w04g],robust cformat(%9,4f) pformat(%5,2f)
sformat(%8,2f)
(sum of wgt is 32,049.6321299076)

Linear regression                               Number of obs     =     32,054
                                                F(1, 32052)       =    6355.04
                                                Prob > F          =     0.0000
                                                R-squared         =     0.5937
                                                Root MSE          =     .39166

------------------------------------------------------------------------------
         |               Robust
  l_kons |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
---------+--------------------------------------------------------------------
  ldochg |     0,7296     0,0092    79,72   0,00       0,7117      0,7476
   _cons |     1,9129     0,0689    27,75   0,00       1,7778      2,0480
------------------------------------------------------------------------------
```