Egzamin z ekonometrii 13.06.2006
II semestr

Pytania teoretyczne

1. Podać ogólną postać modelu DL i ADL i $ARMA(p,q)$

   Model DL
   $$y_t = \mu + \beta_0 x_t + \ldots + \beta_p x_{t-p} + \varepsilon_t$$
   Model ADL
   $$y_t = \alpha_1 y_{t-1} + \ldots + \alpha_p y_{t-p} + \mu + x_t\beta_0 + x_{t-1}\beta_1 + \ldots + x_{t-s}\beta_s + \varepsilon_t$$
   Model $ARMA(p,q)$
   $$y_t = \mu + \alpha_1 y_{t-1} + \alpha_p y_{t-p} + \varepsilon_t + \theta_1\varepsilon_{t-1} + \ldots + \theta_q\varepsilon_{t-q}$$

2. Podaj standardowe założenia MNW dla modelu szacowanego na próbie przekrojowej.

   - znana postać łacznej funkcji gęstości
   - słaba egzogeniczność zmiennych objaśniających względem wektora parametrów szacowanych
   - niezależność obserwacji
   - identyfikowalność parametrów

3. Jakie są wady liniowego modelu prawdopodobieństwa? Odpowiedź uzasadnić.

   - Heteroskedastyczność błędu losowego
   $$\begin{aligned} \mathrm{E}\left(y_i^2\middle|\,\boldsymbol{x}_i\right) &= 0^2\left(1-\boldsymbol{x}_i\boldsymbol{\beta}\right) + 1^2\boldsymbol{x}_i\boldsymbol{\beta} = \boldsymbol{x}_i\boldsymbol{\beta} \\ \mathrm{Var}\left(y_i|\,\boldsymbol{x}_i\right) &= \mathrm{E}\left(y_i^2\middle|\,\boldsymbol{x}_i\right) - \left[\mathrm{E}\left(y_i|\,\boldsymbol{x}_i\right)\right]^2 = \boldsymbol{x}_i\,\boldsymbol{\beta} - \left(\boldsymbol{x}_i\boldsymbol{\beta}\right)^2 \\ &= \boldsymbol{x}_i\boldsymbol{\beta}\left(1-\boldsymbol{x}_i\boldsymbol{\beta}\right) \end{aligned}$$
   $$\mathrm{Var}\left(\varepsilon_i|\,\boldsymbol{x}_i\right) = \mathrm{Var}\left(y_i - \boldsymbol{x}_i\boldsymbol{\beta}|\,\boldsymbol{x}_i\right) = \mathrm{Var}\left(y_i\right) = \boldsymbol{x}_i\boldsymbol{\beta}\left(1-\boldsymbol{x}_i\boldsymbol{\beta}\right)$$
   - Brak gwarancji, że przewidywane prawdopodobieństwo sukcesu $\widehat{y}_i = \boldsymbol{x}_i\,\boldsymbol{b} \in [0,1]$

4. Wyjaśnić, dlaczego za pomocą estymatora efektów stałych nie można oszacować parametrów przy zmiennych nie zmieniających się w czasie.

   Ponieważ przekształcenie efektów stałych eliminuje takie zmienne z szacowanego równania przekształconego. Szacujemy równanie
   $$y_{it} - \overline{y}_i = \left(\boldsymbol{x}_{it} - \overline{\boldsymbol{x}}_i\right)\boldsymbol{\beta} + \left(\varepsilon_{it} - \overline{\varepsilon}_i\right)$$
   ale na przykład, jeśli
   $$\ln\left(\text{płaca}_{it}\right) = \boldsymbol{\beta}_0 + \beta_1\text{płeć}_i + \beta_2\text{wiek}_{it} + u_i + \varepsilon_{it}$$
   to:
   $$\ln\left(\text{płaca}_{it}\right) - \overline{\text{płaca}_i} = \beta_2\left(\text{wiek}_{it} - \overline{\text{wiek}_i}\right) + \varepsilon_{it} - \overline{\varepsilon}_{it}$$
   i w modelu przekształconym skróciła się płeć.

ZADANIE 1 Dany jest proces DL następującej postaci:
$$y_t = \mu + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \theta_0 x_t + \theta_1 x_{t-1} + \varepsilon_t.$$

1. Wyjaśnić, jaka jest interpretacja współczynnika θ_0
2. Dany jest scenariusz bazowy, w którym $x_t = x_t^*$, $x_{t-1} = x_{t-1}^*$, $x_{t-2} = x_{t-2}^*$, $y_{t-2} = y_{t-2}^*$, $y_{t-3} = y_{t-3}^*$
   (a) Policzyć, ile wyniesie $\mathrm{E}\left(y_t\right)$ w scenariuszu bazowym

(b) Policzyć, o ile zmieni się $\mathrm{E}(y_t)$ w stosunku do scenariusza bazowego, jeśli zamiast x_{t-1}^* przyjmiemy $x_{t-1}^{**} = x_{t-1}^* + 2$.

(c) Policzyć, o ile zmieni się $\mathrm{E}(y_t)$ w stosunku do scenariusza bazowego, jeśli przyjmiemy $x_{t-1}^{**} = x_{t-1}^* + 1$, $x_t^{**} = x_t^* - 1$.

3. Ile wynosi mnożnik długookresowy w tym modelu?

4. Jak wygląda równowaga długookresowa w tym modelu?

5. Przyjmijmy, że $\theta_0 = 0$. Jaki dodatkowe ograniczenie musi być prawdziwe, aby x nie był przyczyną w sensie Grangera y?

Rozwiązanie:

1. θ_0 jest mnożnikiem bezpośrednim

(a)

$$\begin{aligned} y_{t-1} &= \mu + \alpha_1 y_{t-2} + \alpha_2 y_{t-3} + \theta_0 x_{t-1} + \theta_1 x_{t-2} + \varepsilon_{t-1} \\ y_t &= \mu + \alpha_1 (\mu + \alpha_1 y_{t-2} + \alpha_2 y_{t-3} + \theta_0 x_{t-1} + \theta_1 x_{t-2} + \varepsilon_{t-1}) \\ &\quad + \alpha_2 y_{t-2} + \theta_0 x_t + \theta_1 x_{t-1} + \varepsilon_t \\ &= \mu (1 + \alpha_1) + \left(\alpha_1^2 + \alpha_2\right) y_{t-2} + \alpha_1 \alpha_2 y_{t-3} + \\ &\quad + \theta_0 x_t + (\alpha_1 \theta_0 + \theta_1) x_{t-1} + \alpha_1 \theta_1 x_{t-2} + \varepsilon_t + \alpha_1 \varepsilon_{t-1} \end{aligned}$$

dla $x_t = x_t^*, x_{t-1} = x_{t-1}^*, x_{t-2} = x_{t-2}^*, y_{t-2} = y_{t-2}^*, y_{t-3} = y_{t-3}^*$

$$\begin{aligned} \mathrm{E}(y_t^*) &= \mu (1 + \alpha_1) + \left(\alpha_1^2 + \alpha_2\right) \mathrm{E}\left(y_{t-2}^*\right) + \alpha_1 \alpha_2 \mathrm{E}\left(y_{t-3}^*\right) \\ &\quad + \theta_0 x_t^* + (\alpha_1 \theta_0 + \theta_1) x_{t-1}^* + \alpha_1 \theta_1 x_{t-2}^* \end{aligned}$$

(b) W tym przypadku

$$\begin{aligned} \mathrm{E}(y_t^{**}) &= \mu (1 + \alpha_1) + \left(\alpha_1^2 + \alpha_2\right) \mathrm{E}\left(y_{t-2}^*\right) + \alpha_1 \alpha_2 \mathrm{E}\left(y_{t-3}^*\right) \\ &\quad + \theta_0 x_t^* + (\alpha_1 \theta_0 + \theta_1) \left(x_{t-1}^* + 2\right) + \alpha_1 \theta_1 x_{t-2}^* \\ &= \mathrm{E}(y_t^*) + 2 (\alpha_1 \theta_0 + \theta_1) \end{aligned}$$

wzrost o $2(\alpha_1 \theta_0 + \theta_1)$

(c) W drugim przypadku mamy

$$\begin{aligned} \mathrm{E}(y_t^{**}) &= \mu (1 + \alpha_1) + \left(\alpha_1^2 + \alpha_1\right) \mathrm{E}\left(y_{t-2}^*\right) + \alpha_1 \alpha_2 \mathrm{E}(y_{t-3}) \\ &\quad + \theta_0 (x_t^* - 1) + (\alpha_1 \theta_0 + \theta_1) \left(x_{t-1}^* + 1\right) + \alpha_1 \theta_1 x_{t-2}^* \\ &= \mathrm{E}(y_t^*) - \theta_0 + \alpha_1 \theta_0 + \theta_1 \end{aligned}$$

a więc wzrost o $-\theta_0 + \alpha_1 \theta_0 + \theta_1$

2. Mnożnik długookresowy w tym modelu jest równy $\theta = \frac{\theta_0 + \theta_1}{1 - \alpha_1 - \alpha_2}$

3. Równowaga długookresowa będzie dana równaniem

$$y^* = \frac{\mu}{1 - \alpha_1 - \alpha_2} + \frac{\theta_0 + \theta_1}{1 - \alpha_1 - \alpha_2} x_t^*$$

4. Dla $\theta_1 = 0$.

ZADANIE 2 Na podstawie danych z badania dochodów i wydatków konsumpcyjnych starano się wyjaśnić wielkość wydatków na zakup nowego samochodu poniesionych przez gospodarstwo domowe w miesiącu badania za pomocą wielkości dochodu gospodarstwa w tym miesiącu. Poniżej znajdują się oszacowania parametrów i efektów cząstkowych dla modelu tobitowego oszacowanego dla tego problemu.

```
Tobit estimates                                   Number of obs   =      35972
                                                  LR chi2(1)      =      36.14
                                                  Prob > chi2     =     0.0000
Log likelihood = -4469.7673                       Pseudo R2       =     0.0040

------------------------------------------------------------------------------
         car |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       dochg |   .4266257    .068442     6.23   0.000     .2924774     .560774
       _cons |  -36639.36    2018.95   -18.15   0.000    -40596.56   -32682.16
-------------+----------------------------------------------------------------
         _se |   14865.12   794.4974           (Ancillary parameter)
------------------------------------------------------------------------------

  Obs. summary:      35679  left-censored observations at car<=0
                       293     uncensored observations
```

```
Marginal Effects: Latent Variable
------------------------------------------------------------------------------
variable |      dF/dx   Std. Err.       z    P>|z|     X_at   [    95% C.I.   ]
---------+--------------------------------------------------------------------
   dochg |   .4266257    .068442     6.23    0.000   1925.31    .292482  .560769
   _cons |  -36639.36    2018.95   -18.15    0.000         1  -40596.4 -32682.3
------------------------------------------------------------------------------

Marginal Effects: Unconditional Expected Value
------------------------------------------------------------------------------
variable |      dF/dx   Std. Err.       z    P>|z|     X_at   [    95% C.I.   ]
---------+--------------------------------------------------------------------
   dochg |   .0034073   .0005466     6.23    0.000   1925.31    .002336  .004479
   _cons |  -292.6207   16.12437   -18.15    0.000         1  -324.224 -261.018
------------------------------------------------------------------------------

Marginal Effects: Conditional on being Uncensored
------------------------------------------------------------------------------
variable |      dF/dx   Std. Err.       z    P>|z|     X_at   [    95% C.I.   ]
---------+--------------------------------------------------------------------
   dochg |   .0396621   .0063628     6.23    0.000   1925.31    .027191  .052133
   _cons |   -3406.25   187.6957   -18.15    0.000         1  -3774.13 -3038.37
------------------------------------------------------------------------------

Marginal Effects: Probability Uncensored
------------------------------------------------------------------------------
variable |      dF/dx   Std. Err.       z    P>|z|     X_at   [    95% C.I.   ]
---------+--------------------------------------------------------------------
   dochg |   6.28e-07   1.01e-07     6.23    0.000   1925.31    4.3e-07  8.3e-07
   _cons |   -.053947   .0029727   -18.15    0.000         1  -.059773 -.048121
------------------------------------------------------------------------------
```

Testy przeprowadzamy na poziomie istotności $\alpha = 0.05$.

1. Wypisać założenia modelu tobitowego.
2. Jaką nietypową cechą będzie się najprawdopodobniej charakteryzować rozkład zmiennej zależnej?

3. Dlaczego policzenie regresji liniowej dla tej całej obserwowanej próby da najprawdopodobniej wartości dopasowane nieinterpretowalne dla części obserwacji?

4. Podać na podstawie wydruku, jaki jest wpływ wzrostu dochodu na prawdopodobieństwo zakupu samochodu i jaki jest wpływ wzrostu dochodu na oczekiwane wydatki na samochód, jeśli konsument zdecydował się go kupić.

5. Podać na podstawie wydruku, jaki jest całkowity wpływ wzrostu dochodu na oczekiwane wydatki na samochód. Czy relacja między znakami tych efektów cząstkowych a oszacowaniami parametrów odpowiada oczekiwaniom?

6. Poniżej podano oszacowania elastyczności dla oczekiwanych wydatków na samochód, policzone dla oszacowanego wyżej modelu policzone dla średniej wartości dochodu w próbie (zatem $\frac{\partial \mathrm{E}(y|x)}{\partial x} \Big/ \frac{y}{x}$ zostało policzone dla $x = \overline{x}$). Zinterpretować te wartości i ocenić, czy są one zgodne z intuicją ekonomiczną.

```
Elasticities after tobit
      y  = E(car*|car>0) (predict, ystar(0,.))
         =  39.293275
------------------------------------------------------------------------------
variable |      ey/ex    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
   dochg |     .16695      .02565    6.51   0.000   .116683  .217217   1925.31
------------------------------------------------------------------------------
```

7. Do zmiennych w modelu dodano zmienną dochód^2 i uzyskano wielkość funkcji wiarogodności na poziomie -4411.0405. Zweryfikować hipotezę, że ta dodatkowa zmienna jest nieistotna w modelu.

   Podpowiedź: $\chi^2_{0.95}(1) = 3.84$, $\chi^2_{0.95}(2) = 5.99$, $\chi^2_{0.95}(3) = 7.81$.

Rozwiązanie:

1. Założenia modelu tobitowego są następujące:

$$y_i^* = \boldsymbol{x}_i \boldsymbol{\beta} + \varepsilon_i$$
$$\varepsilon_i \sim N\left(0, \sigma^2\right)$$
$$y_i = \begin{cases} y_i = y_i^* & \text{dla} \quad y_i > 0 \\ y_i = 0 & \text{dla} \quad y_i \leq 0 \end{cases}$$

   i poszczególne obserwacje są niezależne.

2. Samochód jest dobrem trwałym, więc nie każde gospodarstwo domowe w każdym roku kupuje samochód. W rezultacie próba będzie zawierała dużo zerowych obserwacji. Na podstawie wydruku widzimy, że na 35952 przebadanych gospodarstw, jedynie 293 poniosło wydatki na zakup nowego samochodu.

3. W przypadku oszacowania dla tego modelu zwykłej regresji liniowej najprawdopodobniej część wartości dopasowanych będzie ujemna - co jednak jest bez sensu ponieważ wydatki mogą być wyłącznie dodatnie.

4. Na podstawie wydruku efektów cząstkowych dla prawdopodobieństwa dochodzimy do wniosku, że prawdopodobieństwo zakupu samochodu wzrośnie o 0.0628 punkta procentowego jeśli dochód wzrośnie o 1000 zł. Efekty cząstkowe dla wydatków na zakup samochodu jeśli został on zakupiony odczytujemy z wydruku warunkowych efektów cząstkowych $\mathrm{E}\left(y|\, y > 0\right)$. Wzrost oczekiwanego wydatku na samochód przy 1000 zł wzroście dochodu w grupie, która zakupiła samochód wynosi 39.66 zł.

5. Całkowity wzrost wydatków na samochód odczytujemy z tablicy efektów cząstkowych dla bezwarunkowego y. Wzrost oczekiwanego wydatku na samochód przy 1000 zł wzroście dochodu wynosi 3.4 zł.

6. Oszacowana wielkość oznacza, że przy 1% wzroście dochodu oczekiwane wydatki na samochód wzrosną o $.167\%$. Oszacowanie to wydaje się bardzo niskie wziąwszy pod uwagę, że w Polsce samochód jest ciągle raczej dobrem luksusowym - oszacowanie elastyczności powinno wyjść większe od 1.

7. Statystyka testu LR ma postać: .

$$LR = 2\left(-4411.0405 + 4469.7673\right) = 117.45 > \chi^2_{0.95}(1) = 3.84$$

Testujemy zerowość 1 współczynnika (a więc właściwą wartością krytyczną jest $\chi^2_{0.95}(1)$. Odrzucamy H_0 o nieistotności dochodu^2.

ZADANIE 3 Dany jest model logitowy ze zmiennymi niezależnymi x_1, x_2, x_3. Hipotezy należy testować na poziomie istotności $\alpha = 0.05$.

1. Przetestować łączną nieistotność zmiennych x_1 i x_2. Dla modelu z pełną liczbą zmiennych wielkość logarytmu funkcji wiarogodności wyniosła -277.5. Dla modelu bez zmiennych x_1 i x_2 wartość logarytmu funkcji wiarogodności wyniosła -281. Policzyć odpowiednią statystykę testową i podać wynik testu.

   Podpowiedź: $\chi^2_{0.05}\left(1\right) = 3.84$, $\chi^2_{0.05}\left(2\right) = 5.99$, $\chi^2_{0.05}\left(3\right) = 7.81$.

2. Policzono wielkość statystyki Walda dla hipotezy $H_0 : \beta_1^2 = \beta_2$ oraz $H_0 : \frac{\beta_1^2}{\beta_2} - 1 = 0$ i otrzymano wielkość statystyk testowych 2 i 5. Przeprowadzić testy dla tych dwóch hipotez i skomentować uzyskane wyniki.

3. Dany jest następujący problem: przetestować hipotezę $\gamma = 1$, w modelu

$$y_i = \left(\beta_1 x_{1i} + \beta_2 x_{2i}\right)^\gamma + \varepsilon_i$$

   przy czym $\varepsilon \sim N\left(0, \boldsymbol{I}\right)$. Dystępny program umożliwia szacowanie jedynie zwykłej regresji liniowej. Jakim rodzajem testu należy się posłużyć w tym przypadku?

Rozwiązanie:

1. Statystyka testowa $LR = -2(L_0 - L_1)$ i ma rozkład $\chi^2(k)$, gdzie k to liczba zmiennych o których zakładamy że są nieistotne, L_0 wartość logarytmu wiarogodności dla modelu bez ograniczeń, L_1 wartość logarytmu wiarogodności dla modelu z ograniczeniami. W tym przypadku:

$$LR = -2(-281 - (-277.5)) = 7 > \chi^2(2) = 5.99$$

   Wobec tego należy uznać, że te modele różnią się. Wobec tego zmienne są łącznie istotne.

2. W pierwszym przypadku nie podstaw do odrzucenia $H_0 : \beta_1^2 = \beta_2$ [$2 < 3.84$], w drugim przypadku hipotezę $H_0 : \frac{\beta_1^2}{\beta_2} - 1 = 0$ odrzucamy [$5 > 3.84$]. Analizowane hipotezy są równoważne algebraicznie a więc wynik dwóch testów daje sprzeczne wyniki. W przypadku testu Walda jest to możliwe, ponieważ w małych próbach i dla hipotez nieliniowych wynik tego testu może zależeć od sposobu sformułowania H_0.

3. Statystyką mnożników Lagrange'a - umożliwia on zweryfikowanie H_0 przy znajomości oszacowań modelu z ograniczeniami. W tym przypadku model z ograniczeniami jest bardzo zwykłym modelem liniowym $y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i$ spełniającym założenia $KMRL$.