Instrukcja Obsługi –

Dodatek Linear Regression v 1.1

Jakub Rybacki, jakub.rybacki@gmail.com

Warszawa, 2014

Spis treści

[Słowem wstępu 3](#_Toc382224525)

[Instalacja 3](#_Toc382224526)

[Skróty Klawiszowe 3](#_Toc382224527)

[Wymagania 3](#_Toc382224528)

[Diagnostyka 4](#_Toc382224529)

[Implikowanie zmiennych egzogenicznych 6](#_Toc382224530)

[Detekcja obserwacji nietypowych 7](#_Toc382224531)

[Odfiltrowywanie obserwacji nietypowych 8](#_Toc382224532)

# Słowem wstępu

Niniejszy poradnik powstał w celu przedstawienia działania dodatku Linear\_Regression, oferującego proste i łatwe w użytkowaniu implementacje podstawowych narzędzi diagnostycznych nieobecnych w Analysis Tool Pack. W ramach pakietu użytkownik znajdzie:

* Testy Diagnostyczne (test RESET, Breuscha-Pagana, Durbina-Watsona, Jarque-Bera).
* Detekcje obserwacji nietypowych (Statystyka Cooka i pozostałe miary oparte o statystykę dźwigni)
* Możliwość implikacji zmiennej egzogenicznej na podstawie znanego parametru endogenicznego (rozwinięcie funkcji goal seek dla pełnego zakresu)

Ostatecznie w ramach dodatku dodatkowo załączona została opcja wykrycia współliniowości (Statystyka VIF).

# Instalacja

Instalacja dodatku wprowadza zmiany dwutorowo – w zakładce ‘Dodatki’ pojawi się specjalne menu dla pakietu, takie jak na rysunku poniżej.



Dodatkowo funkcjonalność pakietu zostanie podpisana pod skróty klawiszowe. Te zaprezentowane są poniżej.

## Skróty Klawiszowe

|  |  |
| --- | --- |
| Skrót | Zastosowanie |
| CTRL + SHIFT + F8 | Diagnostyka |
| ALT + SHIFT + F8 | Implikacja zmiennej egzogenicznej |
| CTRL + ALT + F8 | Detekcja obserwacji nietypowych |

## Wymagania

Do poprawnego działania pakietu potrzebna jest instalacja dodatku Solver. Instrukcję obrazkową jak załadować dodatek można znaleźć pod adresem: <http://excel.info.pl/216/jak-wlaczyc-dodate-solver-w-programie-excel.html>

Wszystkie komponenty pakietu zakładać będą ułożenie danych w wierszach.

# Diagnostyka

Wybór opcji ***Basic Diagnostics*** bądź wciśnięcie **CTRL + SHIFT + F8** wywoła pojawienie się następującego interfejsu:



Wybór pól jest stosunkowo intuicyjny – w polu *y* zaznaczamy zakres zmiennych endogenicznych, *x* zmiennych egzogenicznych, *Output* pojedynczą komórkę gdzie wydrukowany ma zostać wynik. Wydruk przeprowadzony na losowo wygenerowanych danych przedstawia się następująco:



Wszystkie statystyki za wyjątkiem testu Jarque-Bera są identyczne w stosunku do wyników z Eviews. W przypadku testu JB różnice są raczej niewielkie, a mają związek z różną definicją wpływu ilości obserwacji/ regresorów (Skośność jak i kurtoza prezentują się identycznie). Prezentowany dodatek liczy wynik za pomocą najbardziej powszechnej formuły:



Czego nie można powiedzieć o pakiecie Eviews.

# Implikowanie zmiennych egzogenicznych

Wybór opcji ***Inverse Function*** bądź wciśnięcie **ALT + SHIFT + F8** wywoła pojawienie się następującego interfejsu:



W polu *Target Cells* zaznaczamy zakres docelowych wartości zmiennej egzogenicznej, które chcemy uzyskać w kolejnych realizacjach funkcji, *Function formula* to miejsce na identyczny zakres pod względem rozmiaru funkcję na podstawie której osiągane są zmienne egzogeniczne, natomiast pole *Changing* powinno zawierać zakres, który chcemy modyfikować. Przykładowe dane przedstawia następujący obrazek:



X wpisane na sztywno, zaznaczone w polu Changing

Wartości docelowe wpisane na sztywno

Formuła (dla 10 wiersza): =1.3\*POTĘGA(H10;3)\*EXP(H10)+5\*H10+2

Przeliczenie funkcji zacznie optymalizacje numeryczną kolumny z formułami, tak aby osiągnęły wartości z lewej kolumny. Należy zauważyć, że makro zakłada, że wszystkie dane znajdują się w pojedynczym arkuszu.

# Detekcja obserwacji nietypowych

Wybór opcji ***Outlier Detection*** bądź wciśnięcie **CTRL + ALT + F8** wywoła pojawienie się następującego interfejsu:



Wybór pól jest analogiczny jak w przypadku diagnostyki – w polu *y* zaznaczamy zakres zmiennych endogenicznych, *x* zmiennych egzogenicznych, *Output* pojedynczą komórkę gdzie wydrukowany ma zostać wynik. Wydruk przeprowadzony na losowo wygenerowanych danych przedstawia się następująco:



Pierwsze 5 kolumn zawiera popularne statystyki związane z obliczaniem wpływu na parametry regresji: statystykę dźwigni, ustandaryzowaną wartość błędu losowego, jej kwadrat, oraz statystyki DFFITS i Cooka. Kolejne pięć pokazuje popularne reguły kciuka, które pozwalają na szybkie odfiltrowanie nietypowych obserwacji.

## Odfiltrowywanie obserwacji nietypowych

Przed rozpoczęciem usuwania obserwacji nietypowych najlepiej skopiować dotychczasowy arkusz (prawy przycisk myszy na nazwie arkusza + przenieś lub kopiuj arkusz). Dla odfiltrowania obserwacji nietypowych konieczne jest umieszczenie odczytu przylegle do danych zmiennych egzogenicznych. Następnie wybieramy ***Dane-> Filtruj*** po czym odfiltrowujemy dane wskazane przez heurystyki.



Następnie wystarczy jedynie skasować podejrzane wiersze.

# Statystyka VIF

Wybór opcji ***Outlier Detection*** wywoła pojawienie się następującego interfejsu:



Wybór pól jest intuicyjny – w polu *x* zaznaczamy zakres egzogenicznych, *Output* pojedynczą komórkę gdzie wydrukowany ma zostać wynik. Wydruk przeprowadzony na losowo wygenerowanych danych gdzie x3=x2-x1 przedstawia się następująco:



Natomiast po wycięciu kłopotliwej danej statystyki przedstawiają się następująco:



Otrzymywane statystyki są zbliżone do wyników Eviews.