Estymacja parametrów modelu liniowego z wieloma zmiennymi objaśniającymi przy pomocy Excela

Liniowy model ekonometryczny o postaci:

$$y_t = \alpha_0 + \alpha_1 X_{1t} + \alpha_2 X_{2t} + \dots + \alpha_k X_{kt} + \xi_t.$$

można oszacować za pomocą wielu pakietów komputerowych.

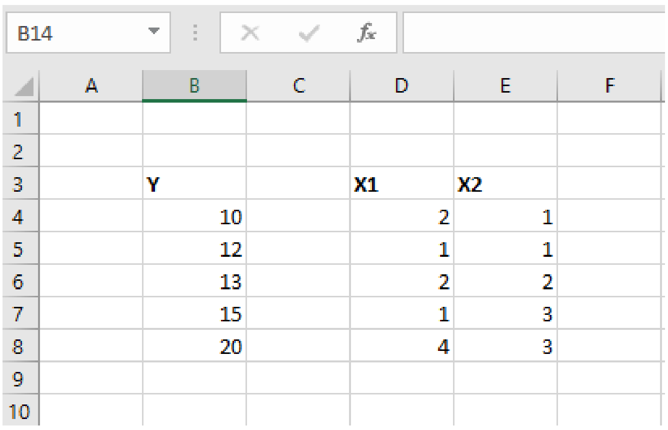
Niniejsze laboratoria mają na celu przybliżenie możliwości szacowania parametrów strukturalnych modelu za pomocą Excela będącego integralną częścią pakietu Office. Wszelkie obliczenia, z pewnością, nie są aż tak zautomatyzowane jak chociażby w pakiecie STATISTICA, Gretl czy też innym, jednak wymagają posiadania pewnych umiejętności (np. posługiwanie się formułami, itp.)
Tym samym dla przykładu przytoczone zostają dane z poprzedniego laboratorium.

Przykład. Na podstawie następujących obserwacji zmiennych *Y*, X_1 i X_2

t	y_t	x_{1t}	x_{2t}
1	10	2	1
2	12	1	1
3	13	2	2
4	15	1	3
5	20	4	3

Oszacować parametry strukturalne modelu liniowego opisującego zależność zmiennej *Y* od zmiennych X_1 i X_2.

W celu oszacowania wspomnianych w poleceniu parametrów a_0, a_1 i a_2 posłużymy się właśnie Excelem.
Na wstępie wprowadzamy dane do arkusza. W całej tej „operacji" mamy swego rodzaju pewną dowolność, jednak dla przejrzystości prezentacji zachowano pewien porządek.



	A	B	C	D	E	F
1						
2						
3		Y		X1	X2	
4		10		2	1	
5		12		1	1	
6		13		2	2	
7		15		1	3	
8		20		4	3	
9						
10						

Estymacji parametrów dokonujemy przy pomocy Klasycznej metody najmniejszych kwadratów (ujęcie macierzowe). Podejście to pozwala na szacowanie parametrów dowolnego modelu liniowego. Tym samym należy wyznaczyć wektor *a*:

$$a = (X^T X)^{-1} \cdot X^T y.$$

Przekształcamy zatem dane w arkuszu tak, aby odpowiadały elementom składowym powyższego działania, czyli wektorowi *y* oraz macierzy *X*. Otrzymujemy:

B	C	D	E	F	G	H	I	J	K	L
Y		**X1**	**X2**			**y**		**X**		
10		2	1			10		1	2	1
12		1	1			12		1	1	1
13		2	2			13		1	2	2
15		1	3			15		1	1	3
20		4	3			20		1	4	3

Należy w tym miejscu przypomnieć, że naszym zadaniem jest oszacowanie parametrów modelu:

$$y_t = \alpha_0 + \alpha_1 X_{1t} + \alpha_2 X_{2t} + \xi_t.$$

stąd też macierz *X* składa się z trzech kolumn. Kwestia ta została już wytłumaczona podczas poprzednich zajęć.
Mając przygotowane w arkuszu macierze składowe możemy rozpocząć obliczenia. Na wstępie wyznaczyć należy iloczyn macierzy X^T oraz *X*. W tym celu należy wstępnie transponować macierz *X*. Wykorzystuje się do tego dostępną formułę **=transponuj(…)** gdzie w nawiasie podaje się zakres komórek macierzy, którą chcemy transponować. Wstępnie jednak należy zaznaczyć w arkuszu zakres komórek w których to macierz transponowana ma się znaleźć. Po zaznaczeniu opisane obszaru, wpisaniu formuły naciskamy kombinację klawiszy **Ctrl+Shift+Enter**. Wówczas uzupełni nam się cały zaznaczony obszar. W przypadku naciśnięcia samego Enter uzupełni nam się tylko pojedyncza komórka arkusza.

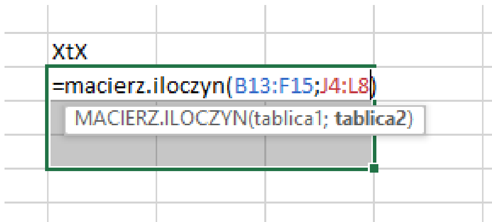
11						
12		Xt				
13		=transponuj(J4:L8)				
14						
15						
16						

A tak wygląda, finalnie transponowana przez nas macierz:

12		Xt				
13		1	1	1	1	1
14		2	1	2	1	4
15		1	1	2	3	3
16						

Mając do dyspozycji macierz X^T oraz *X* możemy pomnożyć te dwie macierze i uzyskać $X^T X$.

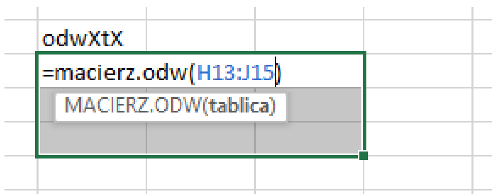
W tym celu zaznaczamy kolejny obszar w arkuszu, w którym to ma znaleźć się oszacowana przez nas macierz X^TX. Ponieważ mnożymy przez siebie macierz w rozmiarze 3x4 przez macierz 4x3 macierz wynikowa będzie miała rozmiar 3x3 (trzy wiersze i 3 kolumny). Aby pomnożyć przez siebie dwie wskazane macierze posługujemy się tym razem formułą: **=macierz.iloczyn(…;…)**, gdzie w nawiasie wpisujemy zakres komórek macierzy X^T a po średniku zakres komórek macierzy X.



W celu wyświetlenia wartości wciskamy ponownie kombinację klawiszy **Ctrl+Shift+Enter**. Wówczas mamy:

XtX		
5	10	10
10	26	22
10	22	24

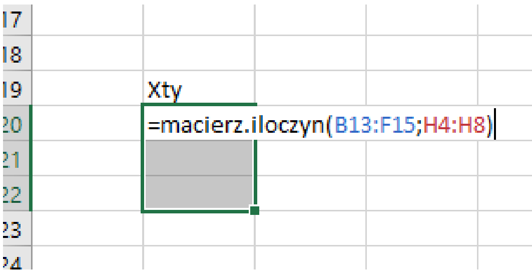
Pozostaje nam oszacowanie macierzy odwrotnej do macierzy X^TX, czyli $(X^TX)^{-1}$. W tym przypadku Excel jest dużo bardziej zautomatyzowany, gdyż posiada wbudowaną formułę pozwalają od razu na odwrócenie macierzy. Przedstawia się ona następująco: **=macierz.odw(…)** , gdzie w nawiasie podajemy zakres komórek arkusza odpowiedzialnych za macierz X^TX. Wygląda to następująco:



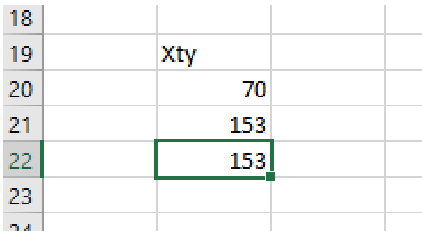
Korzystając z wcześniej opisanego skrótu klawiszowego otrzymujemy:

odwXtX		
1,4	-0,2	-0,4
-0,2	0,2	-0,1
-0,4	-0,1	0,3

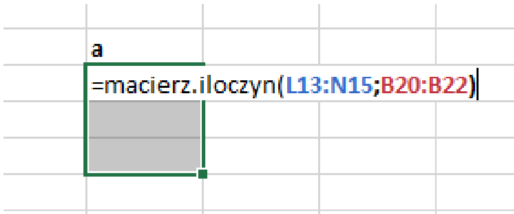
W celu finalizacji szacunków pozostaje nam zatem wyznaczenie macierzy X^Ty. W tym celu mnożymy, wyznaczoną już wcześniej macierz X^T przez wektor y. Wygląda to następująco:



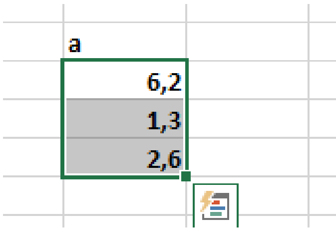
aby otrzymać:



Należy w tym miejscu zauważyć, że w arkuszu zaznaczono obszar 3x1. Wynika to z tego, że mnożyliśmy macierz o 3 wierszach (X^T) przez pacierz o 1 kolumnie (y).
Mając tak oszacowane, elementy składowe możemy wyznaczyć wektor a.



co kończy się wynikiem:



Drugim, znacznie bardziej zautomatyzowanym, sposobem szacowania dowolnych modeli liniowych w pakiecie Excel jest narzędzie **Analysis ToolPak**. Korzystając z pakietu **Analysis ToolPak**, można zaoszczędzić czas i zmniejszyć liczbę czynności wykonywanych podczas opracowywania złożonych analiz statystycznych lub inżynierskich. Użytkownik dostarcza dane i parametry analiz, a narzędzie używa odpowiednich statystycznych lub inżynierskich funkcji makr, aby obliczyć wyniki i wyświetlić je w tabeli wyników. Niektóre narzędzia generują wykresy w dodatkowych tabelach wyników.

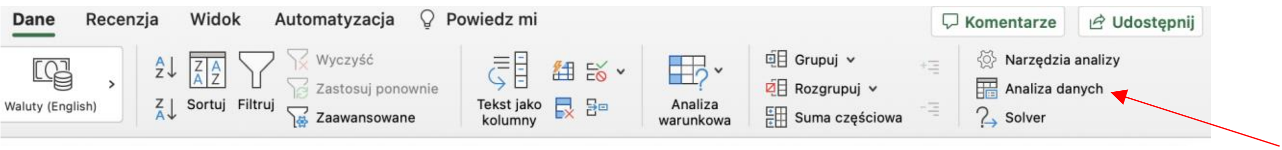
Jeśli korzystamy z Excel 365 (system operacyjny Windows) wspomniane narzędzie dostępne jest w menu **Dane** w zakładce **Analizy** pod nazwą **Analiza danych**.

W przypadku braku narzędzia należy je samemu załadować. W tym celu należy wykonać następujące kroki:

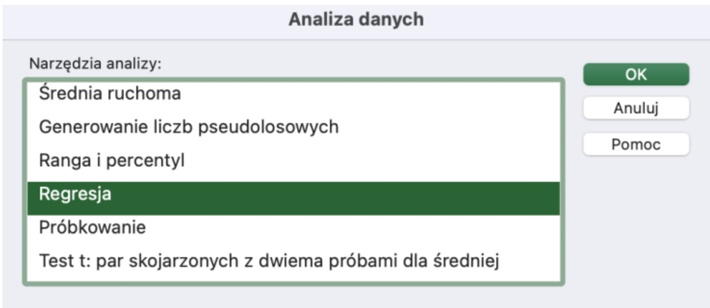
1. Kliknij kartę **Plik**, kliknij pozycję **Opcje**, a następnie kliknij kategorię **Dodatki**.

   Jeśli korzystasz z programu Excel 2007, kliknij przycisk pakietu Microsoft Office a następnie kliknij pozycję Opcje programu Excel.
2. W polu **Zarządzanie** wybierz pozycję **Dodatki programu Excel**, a następnie kliknij przycisk **Przejdź**.
3. W oknie **Dodatki** zaznacz pole wyboru **Analysis ToolPak**, a następnie kliknij przycisk **OK**.
   - Jeśli pozycja **Analysis ToolPak** nie jest wyświetlana w polu **Dostępne dodatki**, kliknij przycisk **Przeglądaj**, aby odnaleźć ten dodatek.
   - Jeśli zostanie wyświetlony monit informujący o tym, że na komputerze nie zainstalowano pakietu Analysis ToolPak, kliknij przycisk **Tak**, aby zainstalować ten pakiet.

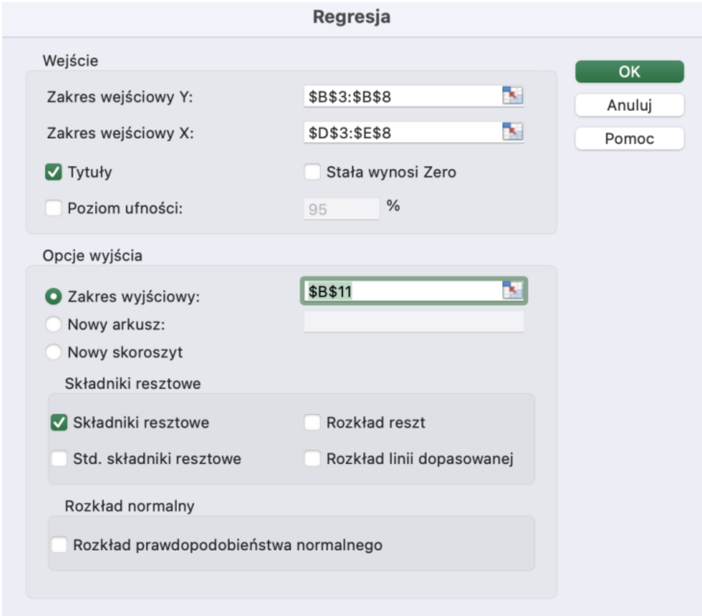
Po wykonaniu tych operacji narzędzie to powinno być dostępne, jak wyżej wspomniano w menu **Dane** w zakładce **Analizy** pod nazwą **Analiza danych**.



Po kliknięciu na Analizę otwiera się okno Narzędzi analizy, które w zależności od wersji Microsoft Office może wyglądać nieco różnie. Dla przykładu:



Modułem dla nas istotnym jest moduł **Regresja**. W celu przeprowadzenia analizy modelu ekonometrycznego (szacowania parametrów strukturalnych, stochastycznych oraz analizy reszt) należy wprowadzić dane wejściowe z arkusza kalkulacyjnego. I tak w analizowanym przez nas przykładzie wygląda to następująco:



Opcję **Tytuły**, wykorzystujemy wówczas, jeżeli wartości zmiennych w arkuszu wskazujemy wraz z ich nazwami. Tak też zostało poczynione w tym przypadku. Opcja **Zakres wyjściowy**, to wskazanie adresu komórki (miejsca w arkuszu), w którym ma zostać wyświetlona analiza. W **Składnikach resztowych** możemy dodatkowo wykonać analizę reszt w różnych przekrojach.
Po takim wyborze opcji otrzymujemy następujące tabele wynikowe:

PODSUMOWANIE - WYJŚCIE

Statystyki regresji	
Wielokrotnoś	0,934953438
R kwadrat	0,874137931
Dopasowany	0,748275862
Błąd standar	1,910497317
Obserwacje	5

ANALIZA WARIANCJI

	df	*SS*	*MS*	*F*	*Istotność F*
Regresja	2	50,7	25,35	6,94520548	0,12586207
Resztkowy	2	7,3	3,65		
Razem	4	58			

	Współczynniki	*Błąd standardowy*	*t Stat*	*Wartość-p*	*Dolne 95%*	*Górne 95%*	*Dolne 95,0%*	*Górne 95,0%*
Przecięcie	6,2	2,260530911	2,74271852	0,11119674	-3,5262795	15,9262795	-3,5262795	15,9262795
X1	1,3	0,854400375	1,52153491	0,26753298	-2,3761881	4,9761881	-2,3761881	4,9761881
X2	2,6	1,046422477	2,48465611	0,13091585	-1,9023925	7,10239253	-1,9023925	7,10239253

SKŁADNIKI RESZTOWE - WYJŚCIE

Obserwacja	*Przewidywane Y*	*Składniki resztowe*
1	11,4	-1,4
2	10,1	1,9
3	14	-1
4	15,3	-0,3
5	19,2	0,8

Otrzymany wynik jest analogiczny do wcześniej otrzymanego. Dodatkowo Excel wyrzuca dużo więcej informacji dotyczących konstruowanego modelu.

Zadania do samodzielnego rozwiązania:

Zadanie 1. **2.14.** Na podstawie następujących obserwacji zmiennych Y, X_1, X_2, X_3:

t	y_t	x_{t1}	x_{t2}	x_{t3}
1	3	−2	0	1
2	3	−1	1	1
3	4	−1	1	2
4	3	1	0	2
5	5	2	1	1
6	4	1	1	2

oszacować parametry strukturalne modelu liniowego opisującego zależność zmiennej Y od zmiennych X_1, X_2, X_3.

Zadanie 2. **2.15.** Na podstawie następujących obserwacji zmiennych Y, X_1, X_2, X_3:

t	y_t	x_{t1}	x_{t2}	x_{t3}
1	3	0	0	2
2	3	1	0	2
3	4	1	1	1
4	5	1	0	1
5	5	2	1	1

oszacować parametry strukturalne modelu liniowego opisującego zależność zmiennej Y od zmiennych X_1, X_2, X_3.