

# Econometric modellings, explained and explaining variables

1 marca 2021



- Jest to nauka pomocnicza stosowana na początku w ekonomii;
- Jest to nauka, która stosuje metody matematyczne, statystyczne i informatyczne dla analiz ilościowych pomagających zrozumienie i wzajemną korelację między dwoma lub więcej zjawiskami społecznymi i przyrodniczymi.

Podstawowym zadaniem ekonometrii jest odpowiedź na różne pytania dotyczące wpływu jednego zjawiska na drugie.

Przykładowo:

- 1 W jaki sposób dawka nawozu mineralnego: azotowego (N), fosforowego (P) i potasowego (K) wpływa na wielkość plonu?
- 2 W jaki sposób populacja pszczoł, trzmieli i innych owadów zapylających wpływa na wielkość plonów?
- 3 W jaki sposób użytkowanie pestycydów zmniejsza populacje pszczoł, trzmieli i innych owadów zapylających?
- 4 W jaki sposób gęstość zaludnienia regionu wpływa na rozprzestrzenianie epidemii COVID-19?
- 5 W jaki sposób rolnictwo, przemysł, turystyka kształtuje produkt krajowy brutto danego kraju?

Odpowiedź na powyższe pytania jest intuicyjnie oczywista, choć nie zawsze:

- 1 Wiadomo, nawóz  $N, P, K$  nie za dużo nie za mało;
- 2 Im więcej pszczoł i trzmieli tym większe plony;
- 3 Im więcej pestycydów tym większe plony, ale też mniej owadów zapylających, więc mniejsze plony. Zatem myśląc o maksymalizacji plonów pestycydów nie może być za dużo;
- 4 Im większa gęstość zaludnienia tym łatwiej rozprzestrzenia się COVID-19? Na początku tak, ale później możliwość transmisji jest ograniczona.
- 5 Wszystkie gałęzie rolnictwo, przemysł i turystyka kształtują produkt krajowy brutto kraju i poprawiają dobrobyt kraju. Ponieważ jednak zasoby są ograniczone, ciężko zadbać o rozwój rolnictwa, przemysłu i turystyki jednocześnie.

Mniej oczywiste pytanie: Czy wzrost wydatków socjalnych wpływa powoduje wzrost dobrobyt społeczeństwa? Z jednej strony:

- 1 Konsumenci mają więcej pieniędzy i więcej kupują;
- 2 Powstają nowe firmy, co wpływa na zmniejsza się bezrobocia;

Z drugiej strony:

- 1 Podwyżki cen, innymi słowy zwiększa się inflacja, a w skrajnych wypadkach brakuje towaru w sklepie;
- 2 Łatwość pozyskania kapitału z innych źródeł niż etat wpływa na decyzje o rezygnacji z mało płatnej pracy;

Zatem decyzja o wydatkach socjalnych jest ryzykowna i nieodwracalna.

- Grupa rolników sprowadza nowy szczep winogrona do Polski w celu produkcji wina;
- W celu opracowania technologii uprawy, pole zostało podzielone na 100 sektorów doświadczalnych;
- Pola podzielono na sektory ze względu na dawki nawozu N,P i K;
- Winogrona sadi się wiosną i owoce zbiera jesienią;
- Jesienią sprawdzamy plony w każdym sektorze i analizujemy plony w zależności od dawki nawozu z danego sektora.

# Example

Niech

- $t = 1, 2, \dots, 100$   
indeks sektora;
- $P_t$  - dawka nawozu  
fosforowego w sektorze  $t$ ;
- $N_t$  - dawka nawozu  
azotowego w sektorze  $t$ ;
- $K_t$  - dawka nawozu  
potasowego w sektorze  $t$ ;
- $G_t$  -  
plony zebrane z sektora  $t$ ;



Celem jest dopasowanie modelu do danych

$$G_t = F_{\beta_0, \beta_1, \dots, \beta_k}(N_t, P_t, K_t, \epsilon_t),$$

gdzie

- $\epsilon_t$  - składnik losowy o zerowej średniej i stałej wariancji;
- $F_{\beta_0, \beta_1, \dots, \beta_k}$  - funkcja zależna od nieznanych parametrów  $\beta_0, \beta_1, \dots, \beta_k$ : funkcja jest zadana intuicyjnie przez użytkownika, a celem jest znalezienie parametrów;
- Zmienne  $N_t$ ,  $P_t$  i  $K_t$  nazywamy **objaśniającymi**;
- Zmienna  $G_t$  nazywa się **objaśniana**;

## Model liniowy:

$$G_t = \beta_0 + \beta_1 N_t + \beta_2 P_t + \beta_3 K_t + \epsilon_t,$$

$\beta_0, \beta_1, \beta_2$  i  $\beta_3$  to nieznane parametry.

## Wada modelu:

- $G_t$  rośnie monotonicznie w  $N_t, P_t$  i  $K_t$  co w praktyce jest niemożliwe ponieważ każdy rolnik wie, że **dawka nawozu nie może być ani za duża ani za mała**;
- gdy  $\beta_k, k = 0, 1, 2, 3$  jest ujemna, istnieje **kombinacja dawek dająca ujemny plon**;

## Model linearyzowalny:

$$G_t = \exp(\beta_0 + \beta_1 N_t + \beta_2 P_t + \beta_3 K_t + \beta_4 N_t^2 + \beta_5 P_t^2 + \beta_6 K_t^2 + \epsilon_t),$$

gdzie  $\beta_k$ ,  $k = 0, 1, \dots, 6$ ,  $\beta_3$  jest nieznanym parametrem. Model może być sprowadzony do liniowego:

$$\ln(G_t) = \beta_0 + \beta_1 N_t + \beta_2 P_t + \beta_3 K_t + \beta_4 N_t^2 + \beta_5 P_t^2 + \beta_6 K_t^2 + \epsilon_t.$$

Teraz

|                      |                                      |
|----------------------|--------------------------------------|
| Zmienna objaśniana   | $\ln(G_t)$                           |
| Zmienna objaśniająca | $N_t, P_t, K_t, N_t^2, P_t^2, K_t^2$ |

Po obliczeniu  $\beta_0, \beta_1, \dots, \beta_6$ :

- Im większa wartość  $G_t$  tym większa wartość  $\ln(G_t)$ ;
- Zatem im większa wartość  $\beta_1 N_t + \beta_4 N_t^2$  tym większa wartość  $\ln(G_t)$ ;
- Podobnie im większa wartość  $\beta_2 P_t + \beta_5 P_t^2$  tym większa wartość  $\ln(G_t)$ ;
- Taka sama sytuacja jest z zależnością  $\beta_3 K_t + \beta_6 K_t^2$  od  $\ln(G_t)$ ;
- Przykładowo, niech  $\beta_0 = 30, \beta_1 = \beta_2 = \beta_3 = 100, \beta_4 = -2, \beta_5 = -1$  and  $\beta_6 = -5$ :

$$\ln(G_t) = 30 + 100N_t + 100P_t + 100K_t - 2N_t^2 - P_t^2 - 5K_t^2 + \epsilon_t.$$

- Taki model daje nam optymalną dawkę:  
 $N_t = 25, P_t = 50, K_t = 10$ ;

Wada modelu:

- Plony  
zależą od dawki nawozu,  
ale nie zależą od proporcji  
między nawozami:  
optymalna dawka  
azotu to 25 niezależnie  
od dawki potasu.  
Taka sytuacja nie może  
mieć miejsca w praktyce.
- W praktyce, zmiana dawki  
azotu wymusza zmiany pozostałych dawek;



Możliwe są inne dopasowania modelu:

$$\begin{aligned}\ln(G_t) = & \beta_0 + \beta_1 N_t + \beta_2 P_t + \beta_3 K_t \\ & + \beta_4 N_t^2 + \beta_5 P_t^2 + \beta_6 K_t^2 \\ & + \beta_7 N_t P_t + \beta_8 P_t K_t + \beta_9 N_t K_t + \epsilon_t.\end{aligned}$$

Można wziąć pod uwagę również parametry gleby: pH - poziom zakwaszenia, wilgotność, poziom zanieczyszczeń etc.

- Modele ekonometryczne jest stale poprawiane aby być bardziej wiarygodny;
- Modele powinny odzwierciedlać zjawiska znane z dotychczasowych doświadczeń i obserwacji;
- Mając odpowiedni model, celem jest dopasowanie parametrów do danych;
- Kolejnym krokiem jest analiza diagnostyczna modelu: czy pasuje do danych, czy model pasuje do założeń, czy przeczy założeniom.
- Ostatecznym celem jest opracowanie prognoz.