

Estymacja bayesowska modelu objaśniającego wolumen sprzedaży ogórków w USA

Dane

Dane pochodzą ze strony amerykańskiego departamentu rolnictwa i zawierają różne informacje (m. in. wolumen i ceny stałe) na temat szerokiej gamy produktów rolnych¹. Plik zawiera jednak o wiele więcej informacji, dlatego plik z danymi użytymi do estymacji jest załączony do raportu.

Użyte zmienne

Zmienne wykorzystane w modelu to:

- Wolumen sprzedanych ogórków (miliony funtów) – zmienna objaśniana,
- Cena za 100 funtów ogórków (hundredweight),
- Cena za 100 funtów pomidorów,
- Cena za 100 funtów cebuli.

Ceny podane są w cenach stałych, gdzie indeks wynosi 100 dla roku 2012. Dane raportowane są dla poszczególnych lat, przy czym zakres próby obejmuje lata 1970-2021.

Finalnie daje to 52 obserwacje i 3 zmienne objaśniające.

Warzywa zostały dobrane ze względu na dostępną wiedzę a priori, dodatkowo pod uwagę brana była popularność produktu. Do analizy nie można było włączyć wszystkich warzyw, ze względu na dość małą liczbę obserwacji (13 obserwacji na szacowany parametr).

Tło

Od lat 60. do wczesnych lat 90. w Stanach Zjednoczonych prowadzone były badania dotyczące elastyczności różnych produktów rolnych²³ na podstawie danych rocznych. Dostarczają one wiadomości, jak kształtowała się w tamtym okresie elastyczność własna i krzyżowa całej gamy produktów rolnych. Dorobek z tych badań zostanie użyty w dalszej części pracy do określenia parametrów a priori.

Transformacje zmiennych

Na wszystkich zmiennych przeprowadzona została transformacja logarytmiczna, co umożliwia interpretację parametrów w ramach elastyczności i jednocześnie wprowadzenie wiedzy a priori do modelu.

¹ <https://www.ers.usda.gov/data-products/vegetables-and-pulses-data/vegetables-and-pulses-yearbook-tables/>, dostęp: 19.04.2023

² Henneberry, S.R., K.P. Piewthongngam, H. Qiang. *Consumer Safety Concerns and Fresh Produce Consumption*, Journal of Agricultural Resource Economics 24 1999

³ You, Z., J.E. Epperson, C.L. Huang. *A Composite System Demand Analysis for Fresh Fruit and Vegetables in the United States*, Journal of Food Distribution Research 1996

Elicytacja parametrów a priori

Strona departamentu rolnictwa umożliwia dostęp do oszacowań elastyczności w formie przejrzystej tabeli⁴, to z niej zostały zaczerpnięte poniższe parametry a priori:

Towar	Rodzaj elastyczności	Oszacowanie
Ogórki	Własna	-0,57
Pomidory	Krzyżowa	-0,04
Cebula	Krzyżowa	-0,11

Dla stałej w modelu niestety nie można zdobyć wiedzy a priori, w związku z czym dla tego elementu zostanie ustalona bardzo duża wariancja, a samo oszacowanie wyniesie 10.

Badania były robione w latach 1970-1992 i 1960-1993, czyli lata się na siebie nakładają. Dodatkowo po zgłębieniu prac okazuje się, że dane o cenach były zbierane z kilku źródeł i w paru latach nie było danych, niech więc okres próby zostanie ustalony na 30 obserwacji (najdłuższe badanie zaokrąglone do pełnych dziesiątek).

Niech reszta w wysokości jednego odchylenia standardowego oznacza błąd predykcji wynoszący 40% (model na logarytmach). Zatem wartość \underline{s}^{-2} wynosi 16.

Na stronie departamentu rolnictwa znajduje się rubryka „odchylenie standardowe”, jednak jest ono liczone z 2 dostępnych badań, dlatego informacja ta została zignorowana. Wymusza to ustalenie odchyleń standardowych, niech więc:

- ogórki raczej nie są dobrem luksusowym, więc przedział nie powinien zawierać wartości większych od 0. Niech więc przedział (-1.07, -0.07) zawiera 95% masy gęstości
- oszacowania elastyczności wskazywałyby, że pomidory i cebula są dobrami komplementarnymi, jednak sensowne wydaje się również patrzenie na te produkty jak na substytuty (w szczególności pomidory). W związku z tym, a także z faktem, że oszacowania empiryczne są bliskie 0, niech więc 95% masy gęstości zawiera się w przedziale (-0.7, 0.7) (odpowiednio przesuniętym dla produktu)

Dla powyższych założeń:

- odchylenie standardowe stałej jest duże
- odchylenie standardowe elastyczności własnej ogórków wynosi około 0,25 (wariancja 0,0625)
- odchylenie standardowe elastyczności pomidorów i cebuli wynosi około 0,36 (wariancja 0,1296)

Informacje te pozwalają na określenie parametrów a priori:

- \underline{s}^{-2} wynosi 6.25
- liczba obserwacji „poprzedniego” modelu wynosi 30
- oszacowanie parametrów wynoszą odpowiednio: 10, -0.57, -0.04, -0.11 (stała, ogórki, pomidory i cebula)

⁴ https://data.ers.usda.gov/reports.aspx?ID=17825#P3809bf1ae6a54082a55a1a41eb7ac5f9_4_214, dostęp: 20.04.2023

- elementy diagonalne macierzy U wynoszą: 200, 0.39, 3.8, 3.8, poza diagonalną ze względu na brak danych elementy wynoszą 0.

Wyniki estymacji

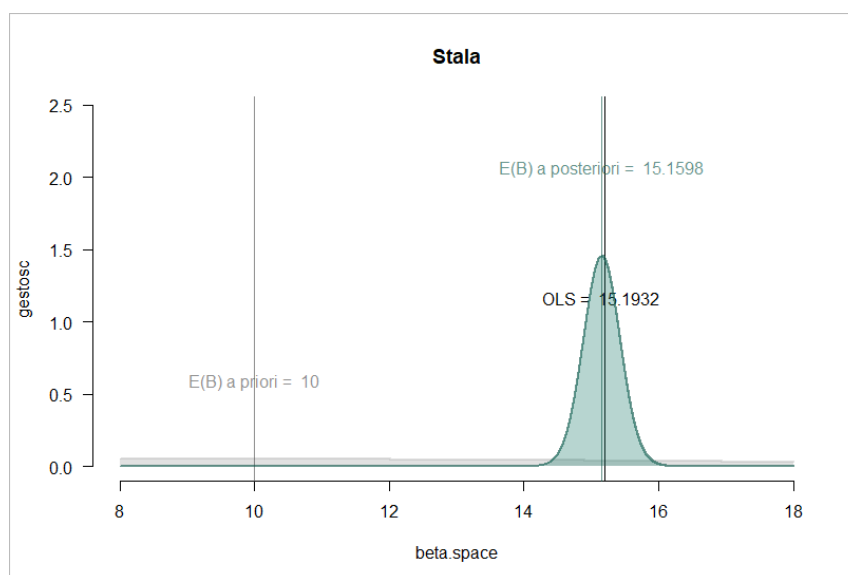
Wartości oczekiwane a posteriori

Wartości oczekiwane parametrów a posteriori prezentują się następująco:

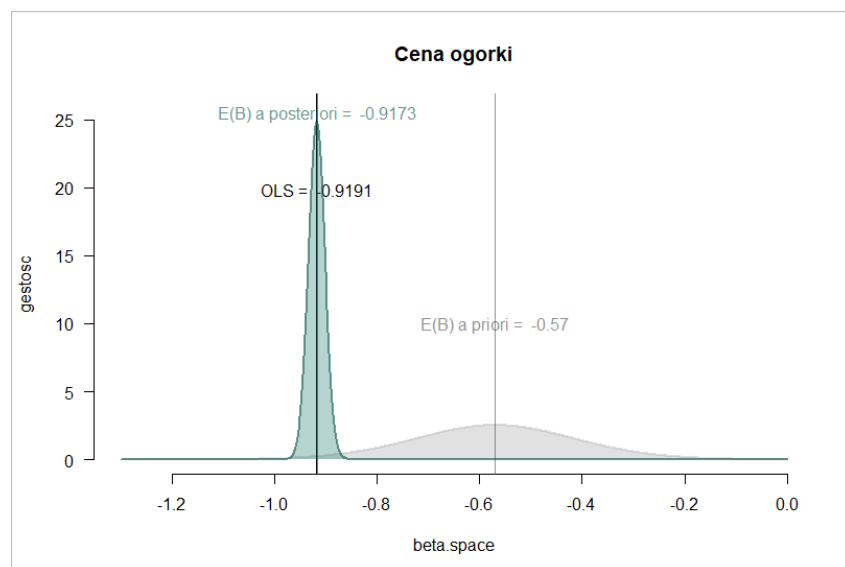
Stała	15,160
Cena ogórków	-0,917
Cena pomidorów	-0,812
Cena cebuli	-0,630

Rozkłady brzegowe parametrów a posteriori

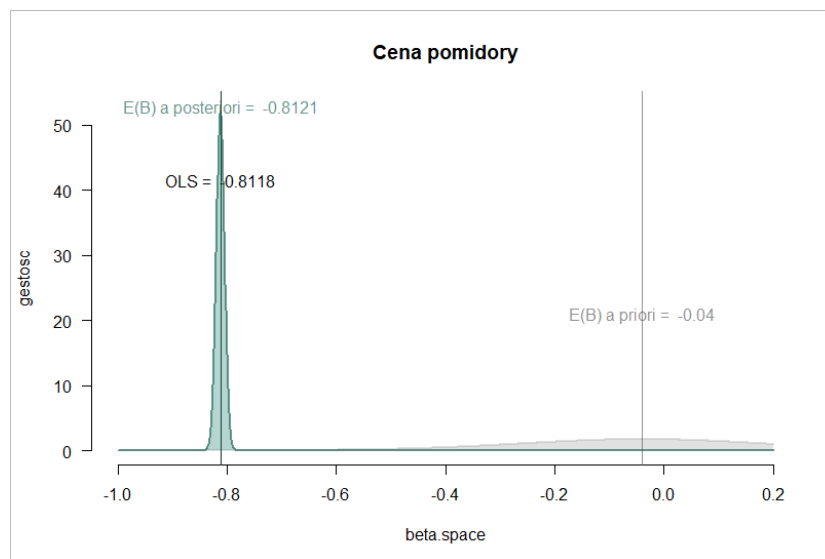
Stała:



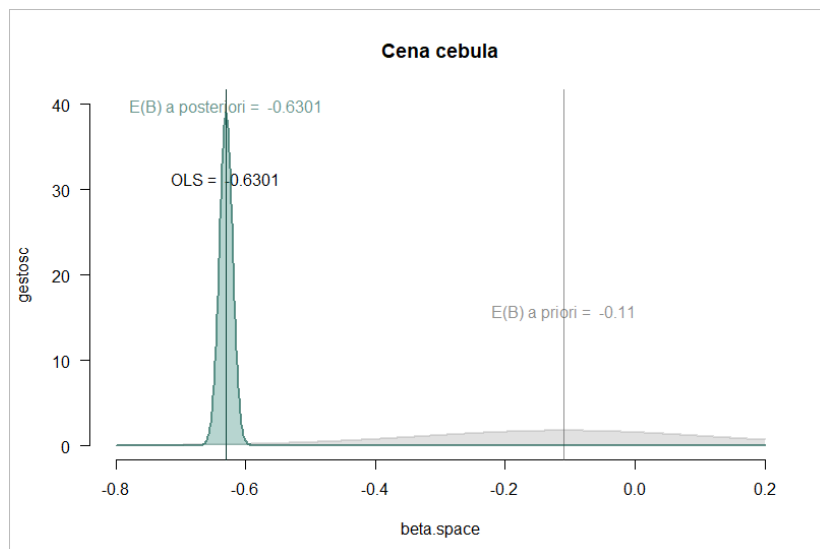
Cena ogórków:



Cena pomidorów:



Cena cebuli:



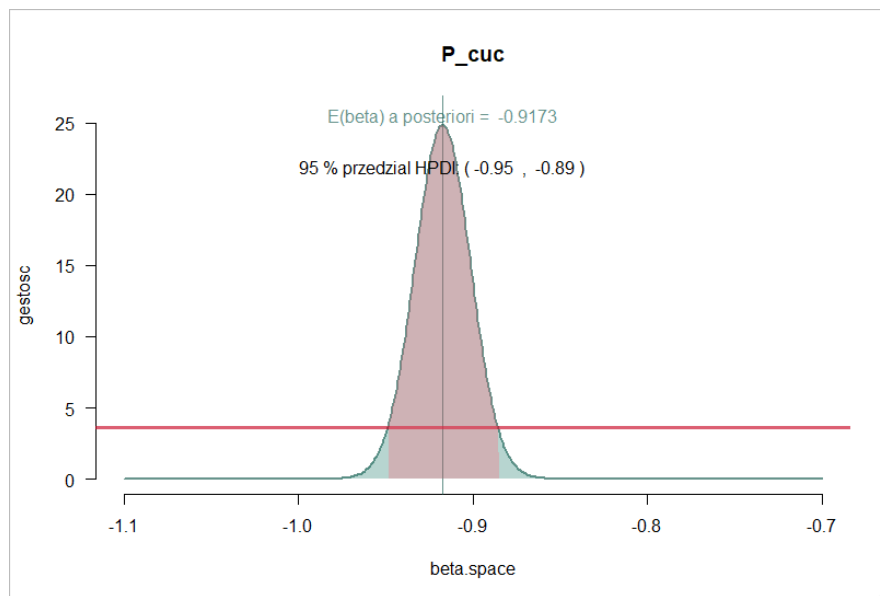
Należy zwrócić uwagę na różne skale wykresów.

Można zauważyć, że dane z modelu zdecydowanie dominują wiedzę a priori. Warto jednak zauważyć, że od zakończenia badań minęło 30 lat, także nie jest to w żaden sposób zaskakujące, parametry a priori zostały ustalone w sposób to uwzględniający. Dodatkowo model KMNK ma dopasowany R^2 wynoszący około 70%, co jest dość wysokim wynikiem. Ma to odzwierciedlenie w małej wariancji oszacowanych parametrów i „dużej sile przyciągania”.

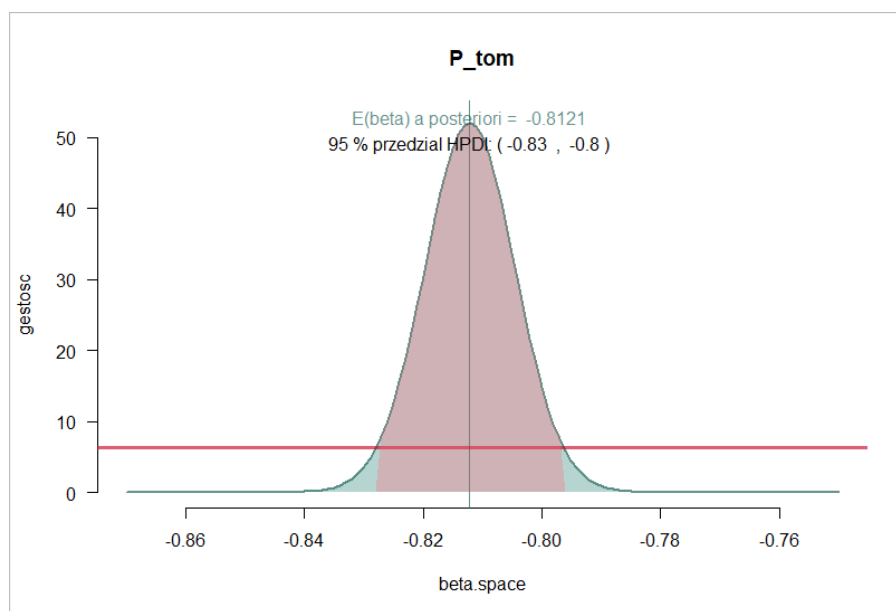
Znaczenie zmiennych

HPDI

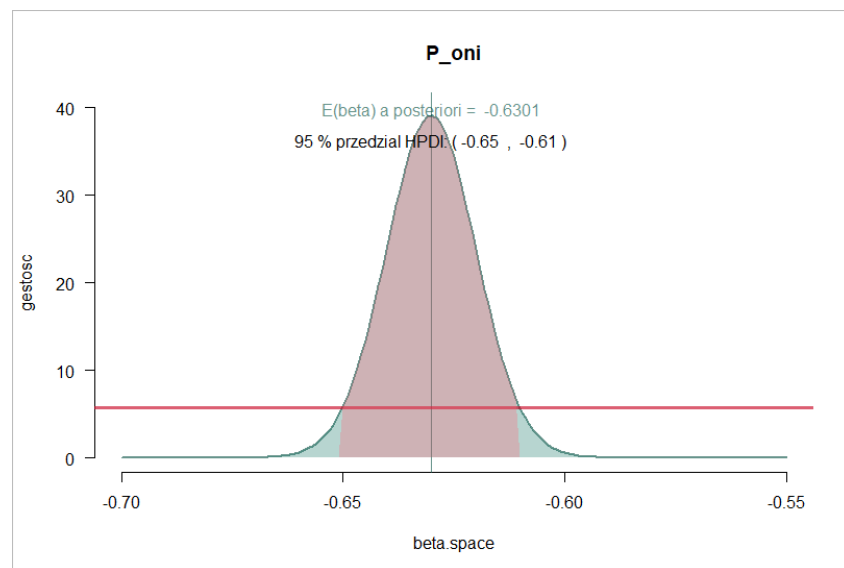
HDPI dla parametru przy zmiennej cena ogórków:



HDPI dla parametru przy zmiennej cena pomidorów:



HDPI dla parametru przy zmiennej cena cebuli:



Czynniki Bayesa

Czynniki Bayesa dla poszczególnych zmiennych prezentują się następująco:

Zmienna	Czynnik Bayesa liczony analitycznie
Cena ogórków	107,0444
Cena pomidorów	824,4273
Cena cebuli	32211,7028

Podsumowanie znaczenia zmiennych

Czynniki Bayesa dla 2 ostatnich zmiennych niezależnie od źródła (>200) mogą zostać zakwalifikowane jako istotne ponad wszelką wątpliwość. Zmienna „Cena ogórków” też powinna zostać umieszczona w modelu, jednak jej „moc” jest mniejsza od poprzedniczek. Przedziały HDPI są tutaj zgodne z czynnikami Bayesa, maksymalny 95% przedział HDPI ma szerokość 0.06, co przy oczekiwanych wartościach parametrów mniejszych niż -0.5 eliminuje wszelkie prawdopodobieństwo, że którykolwiek z parametrów wynosi 0.

Prognozy

Do prognozy została wybrana 20 obserwacja, która prezentuje się następująco:

Zmienna	Wolumen ogórki	Cena ogórki	Cena pomidory	Cena cebula
Wartość	7,14	3,32	3,99	2,92

Prognoza z modelu dla 90% przedziału ufności wygląda następująco:

Dolny próg	Prognoza punktowa	Górny próg
6,42	7,03	7,65

Prognoza z modelu dla 50% przedziału ufności wygląda następująco:

Dolny próg	Prognoza punktowa	Górny próg
6,78	7,03	7,28

Dla obu przedziałów ufności prawdziwa wartość zmiennej się w nich zawiera. Na podstawie jednej obserwacji trudno się wypowiadać o zdolności predykcyjnej modelu, ale wynik jest obiecujący.