

Sprawozdanie DA

Zależność ceny gazu od prądu w Unii Europejskiej

Wykonali:

- Jan Sawicki
- Mateusz Sołtys

Repozytorium projektu:

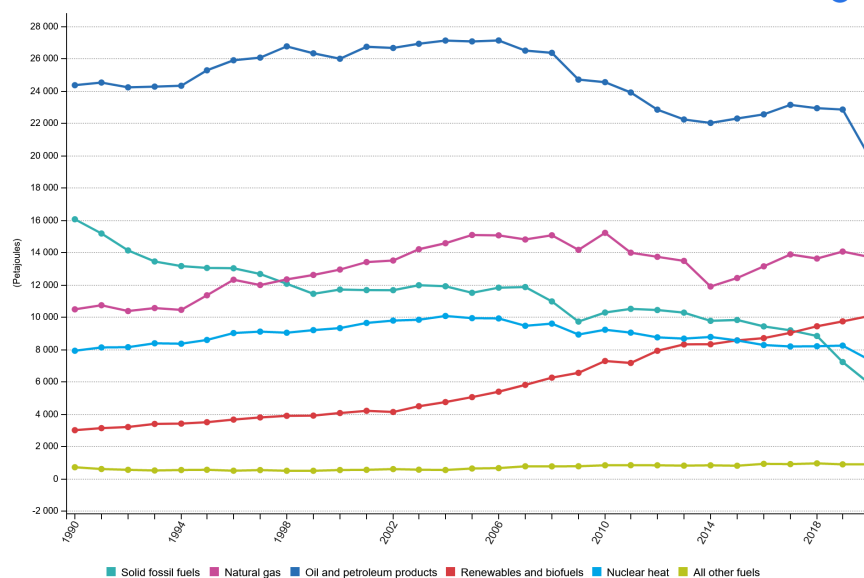
github.com/JTSawicki/Projekt-DataAnalytics-2022

1. Krótkie wprowadzenie do zagadnienia

Ceny energii elektrycznej w większości krajów są bezpośrednio zależne od cen paliw i surowców wykorzystywanych do produkcji energii. Przykładowo w Unii Europejskiej

większość energii elektrycznej jest produkowana z wykorzystaniem paliw kopalnych w szczególności gazu ziemnego i ropy importowanych z krajów trzecich takich jak Rosja czy kraje Zatoki Perskiej. Powoduje to silną zależność cen od surowców importowanych.

Gross available energy, EU, 1990-2020



Rys. 01 Źródła energii w Unii Europejskiej

Analiza zależności tych cen pozwoliłaby na stworzenie modelu pozwalającego przewidzieć szacowany wzrost cen energii elektrycznej w przypadku wzrostu ceny paliwa np. w wyniku embarg i wojny na Ukrainie.

2. Źródło danych

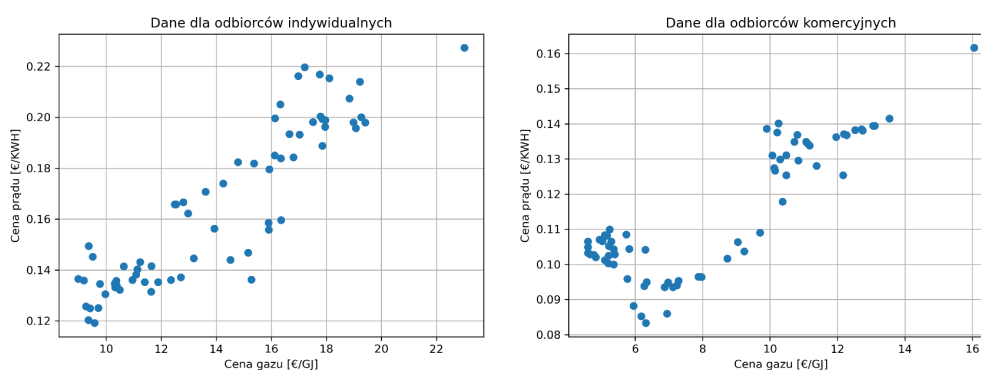
Podczas pozyskiwania danych i ich późniejszej analizy Europa została potraktowana jako byt jednolity w ujęciu statystycznym. Można to zrobić ze względu na silne połączenie interkonektorami pomiędzy poszczególnymi krajami, które sprawia że powstaje de facto jedna sieć elektryczna oraz na "globalność" cen surowców energetycznych. Dlatego wykorzystano jako źródło danych uśrednione ceny energii pozyskane ze strony eurostatu.

Dostępne były tam jedynie dane co do cen prądu i gazu jednak stwierdzono że można na tym oprzeć model ponieważ ceny pozostałych paliw są bezpośrednio zależne od siebie wzajemnie.

Jako jednostkę danych wybrano z dostępnych opcji MJ dla gazu ziemnego i kWh dla energii elektrycznej. Natomiast wszystkie ceny są w euro(€).

Następnie uśredniono dane dostępne dla 25 krajów UE które posiadały wymagane dla modelu informacje.

Dodatkowo dane są podzielone osobno na klientów indywidualnych i komercyjnych ponieważ w przypadku gwałtownych wzrostów cen większość rządów w UE przyznaje pakiety pomocowe(np. obniżając opodatkowanie lub opłaty klimatyczne) dla gospodarstw domowych.



Rys. 02 Pozyskane dane

Kraje dla których uśredniano dane: Austria, Belgia, Bułgaria, Czechy, Niemcy, Dania, Estonia, Hiszpania, Francja, Chorwacja, Węgry, Irlandia, Włochy, Litwa, Luksemburg, Łotwa, Niderlandy, Polska, Portugalia, Rumunia, Szwecja, Słowacja, Turcja, Wielka Brytania.

3. Model i priory

Zdecydowano się na dwa podstawowe modele. Pierwszym z nich jest model liniowy i opisuje on cenę wynikającą z kosztów produkcji energii:

$$y = \alpha + \beta * x$$

Gdzie w modelu:

- y - cena energii za 1 KWh
- x - cena gazu za 1 MJ
- α (alpha) - koszty stałe
- β (beta) - współczynnik konwersji

Jako założenia wstępne modelu(prior) ograniczono wartości α i β do wartości dodatnich ponieważ ze względu na praktyczne ekonomiczne zależności niemożliwym jest ujemne sprzężenie zwrotne przy produkcji od ceny surowca. Dodatkowo silnie ograniczono wartość rozkładu normalnego β do $\text{normal}(0.007, 0.01)$. Wynika to z liniowej zależności konwersji MJ gazu na KWh prądu:

$$\begin{aligned} \text{Prąd}[Kwh] &= \frac{\text{Gaz}[MJ]}{\text{ŚredniaSprawność} * \text{PrzelicznikMJnaKWh}} \\ 1 \text{ Prąd}[Kwh] &= \frac{1 \text{ Gaz}[MJ]}{\frac{(0.45 + 0.57)}{2} * 277.78} \approx 0.0070 * 1 \text{ Gaz}[MJ] \end{aligned}$$

Jednak ponieważ na cenę energii mają również inne czynniki takie jak na przykład subsydia państwowe dla obywateli przy dużych cenach czy to że niektóre sektory gospodarki ograniczą zużycie, a inne zapłacą każdą cenę za potrzebną energię nie ograniczając zużycia, zdecydowano się na stworzenie drugiego modelu opisującego zagadnienie wielomianem czwartego stopnia.

$$y = \alpha + \beta x + \gamma x^2 + \delta x^3 + \epsilon x^4$$

Na α i β przyjęto takie same ograniczenia jak wcześniej natomiast kolejne współczynniki są również nieujemne ograniczone przez rozkład $\text{normal}(0, 0.001)$. Jest tak ponieważ w badanym okresie czasu nie było w Europie dramatycznych katastrof energetycznych(np. całkowity brak dostaw paliwa), a normalne dodatkowe zjawiska rynkowe nie mogą mieć większego wpływu na cenę energii niż produkcja.

Dodatkowo w celu lepszego pokazania zależności obliczenia wykonano również dla dodatkowych modeli wielomianowych drugiego i trzeciego stopnia.

4. Analiza wyników modeli

Kompilacja i dopasowanie danych do modeli trwa strasznie długo. Na wykorzystaniu platformie z sześciordzeniowym procesorem i5-8600K podkręconym do 4,2 GHz trwało to za każdym razem około 1,5 minuty 😞.

Po wykonaniu dopasowań modeli uzyskano następujące rezultaty. Sigma jest wyliczonym odchyleniem standardowym danych.

Dla klientów indywidualnych

Model liniowy:

$$\alpha = 5.17222 \cdot 10^{-2}$$

$$\beta = 8.04387 \cdot 10^{-3}$$

$$\sigma = 0.0135775 \cdot 10^{-}$$

Całkowita różnica rozkładu prawdopodobieństwa danych wejściowych i wyjściowych = 459.789

Model 4 stopnia:

$$\alpha = 8.32107 \cdot 10^{-2}$$

$$\beta = 4.06172 \cdot 10^{-3}$$

$$\gamma = 7.56164 \cdot 10^{-5}$$

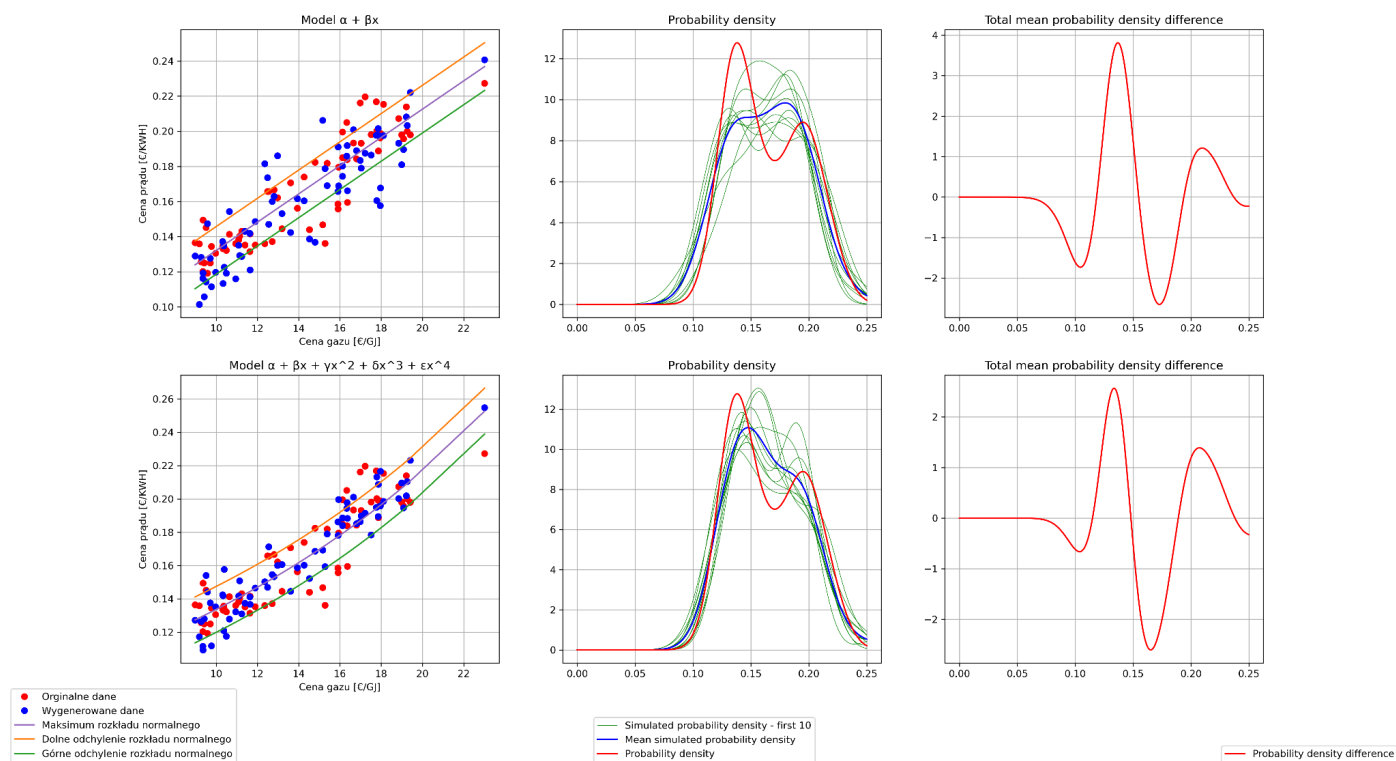
$$\delta = 1.77194 \cdot 10^{-6}$$

$$\epsilon = 5.12730 \cdot 10^{-8}$$

$$\sigma = 0.0137871 \cdot 10^{-}$$

Całkowita różnica rozkładu prawdopodobieństwa danych wejściowych i wyjściowych = 355.638

Klienci indywidualni



Dla klientów komercyjnych

Model liniowy:

$$\alpha = 7.06123 \cdot 10^{-2}$$

$$\beta = 5.25652 \cdot 10^{-3}$$

$$\sigma = 0.00996989 \cdot 10^{-1}$$

Całkowita różnica rozkładu prawdopodobieństwa danych wejściowych i wyjściowych = 449.915

Model 4 stopnia:

$$\alpha = 8.84904 \cdot 10^{-2}$$

$$\beta = 1.37348 \cdot 10^{-3}$$

$$\gamma = 1.08631 \cdot 10^{-4}$$

$$\delta = 5.12428 \cdot 10^{-6}$$

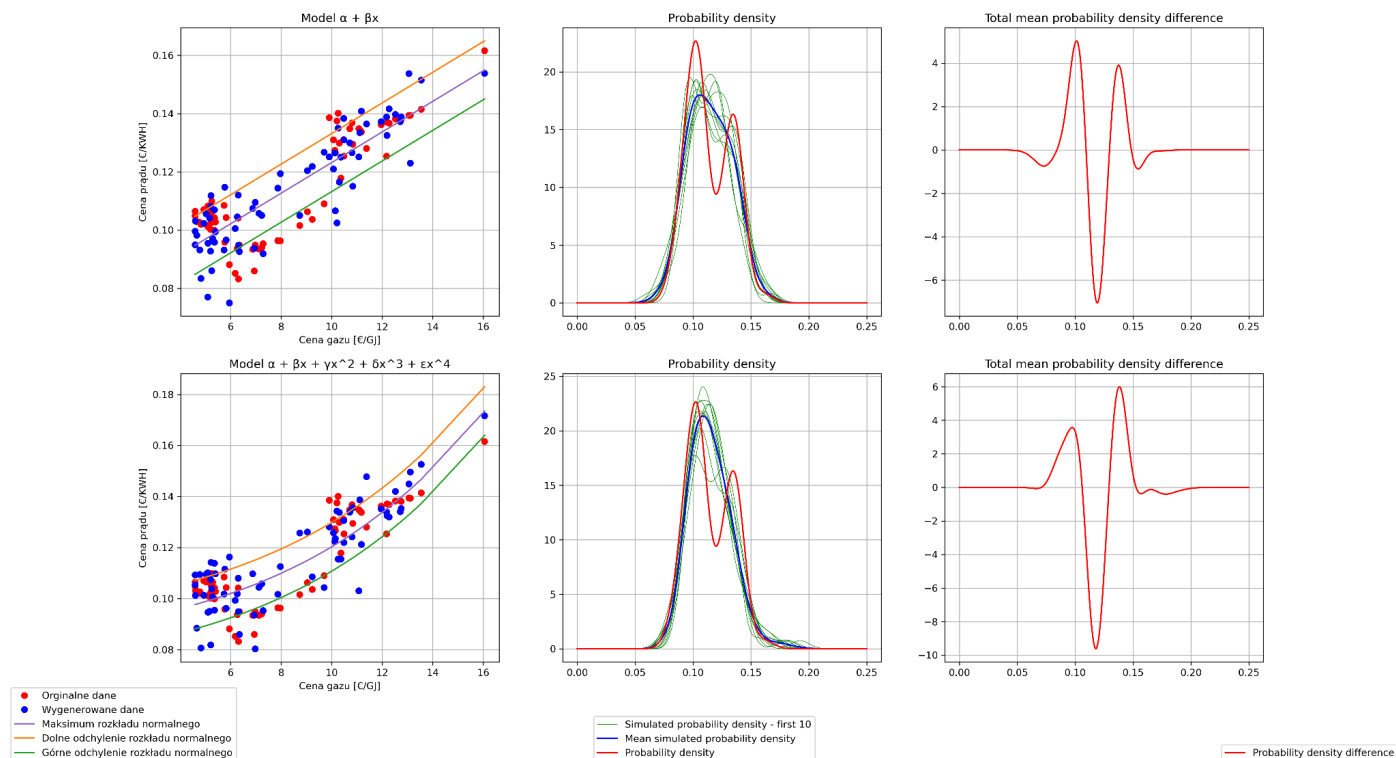
$$\epsilon = 2.07216 \cdot 10^{-7}$$

$$\sigma = 0.00952928 \cdot 10^{-1}$$

Całkowita różnica rozkładu prawdopodobieństwa danych wejściowych i wyjściowych = 591.679

Energy statistics - an overview

Klienci komercyjni



Dla wszystkich rezultatów(również modeli 2 i 3 stopnia) funkcja diagnostyki dopasowania stan zwracała informację o braku problemów.

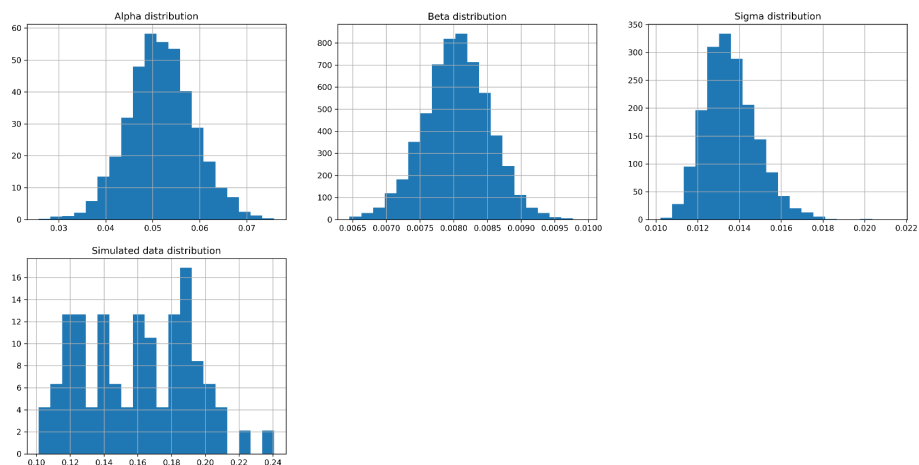
Jak widać z powyższych przykładów zależność liniowa jest dobrą estymatą, a w przypadku klientów komercyjnych wręcz najlepszą.

W przypadku dopasowania do danych gospodarstw domowych były osiągnane nieznacznie lepsze rezultaty w przypadku stosowania modeli wyższego rzędu jednak można postawić pytanie czy zwiększanie złożoności modelu jest usprawiedliwione wobec niewielkiej poprawy.

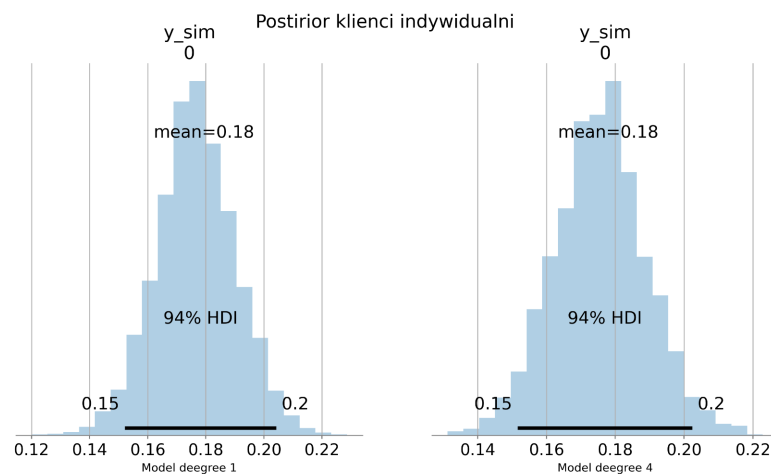
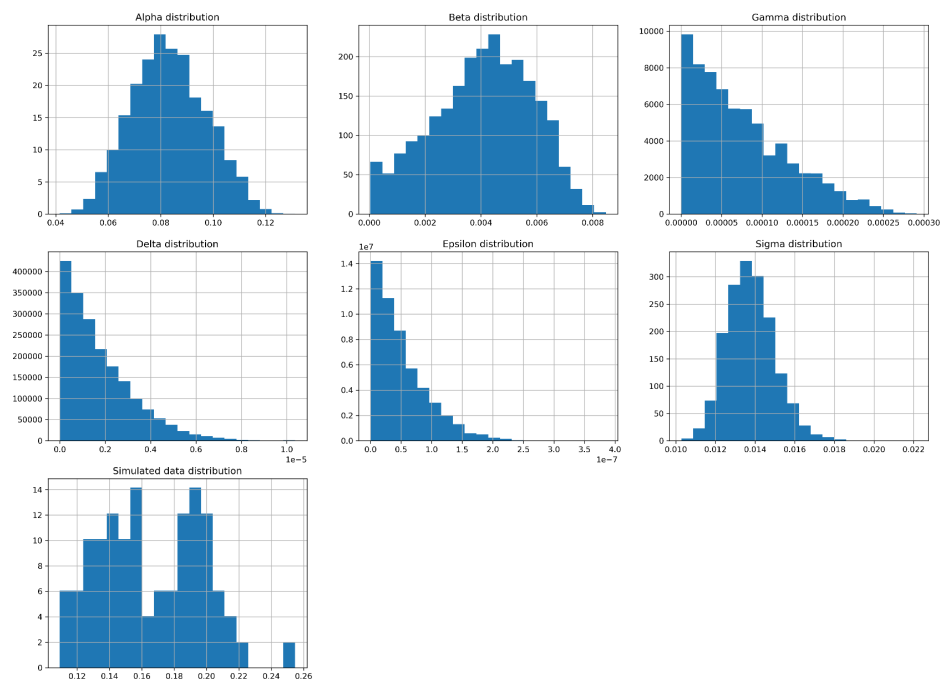
5. Histogramy parametrów & posterior

Klienci indywidualni:

Klienci indywidualni histogramy - model liniowy

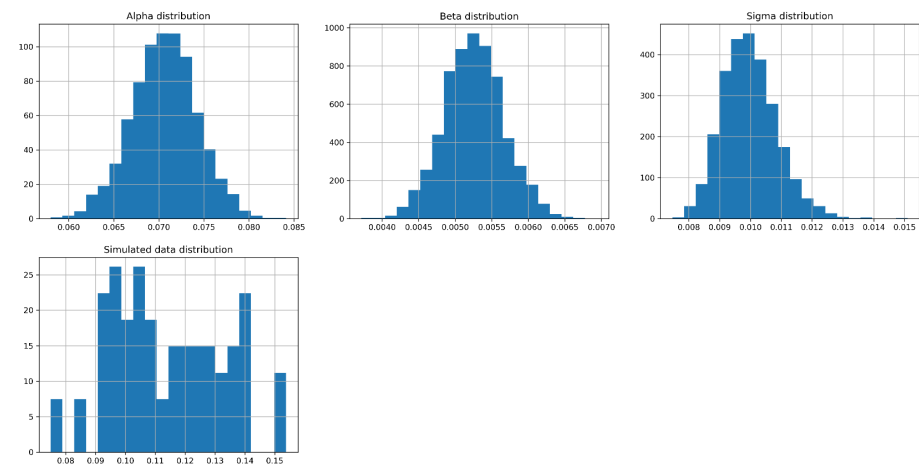


Klienci indywidualni histogramy - model 4 stopnia

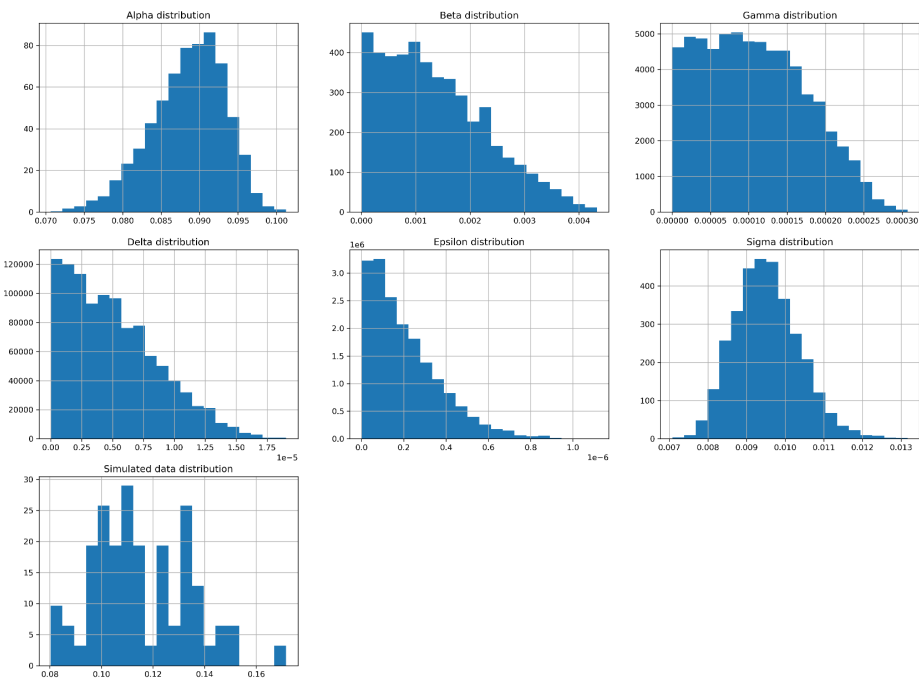


Klienci komercyjni:

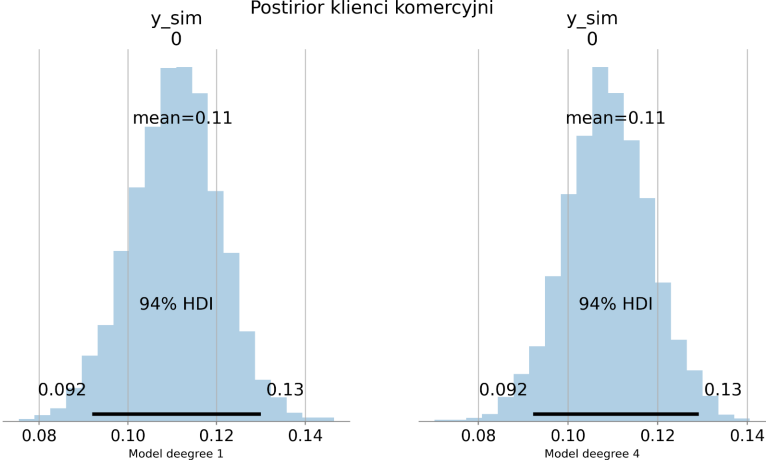
Klienci komercyjni histogramy - model liniowy



Klienci komercyjni histogramy - model 4 stopnia



Posterior klienci komercyjni



6. Bibliografia

- [Eurostat energy statistics - an overview](#)
- [Eurostat Database](#)
- [ScienceDirect](#)