Obraz zawierający zrzut ekranu, Grafika, logo, projekt graficzny

Opis wygenerowany automatycznie

# Projekt 6 *Weryfikacja i interpretacja modelów*

Autorzy:

Kacper Prorok

Popkiewicz Szymon

Celem naszego projektu było oszacowanie modelu, który na podstawie zmiennych objaśniających takich jak: gęstość zaludnienia, ludność powiatu, przystanki na 10 tys. mieszkańców, odsetek ludności powyżej 65 lat, średniej ceny za metr kwadratowy, liczba parków na 10 tys. mieszkańców, średni dochód oszacuje liczbę samochodów o napędzie zielonym na 10 000 mieszkańców.

*W projekcie przyjęliśmy alfa=0.05 dla wszystkich przeprowadzonych testów.*

# Model dla powiatów grodzkich

## Poprawność modelu według założeń (Gaussa-Markowa)

|  |  |
| --- | --- |
| Test | p-value |
| Gold.Quand | 0,1387 |
| Reset-test | 0,8722356 |
| Shapiro-test | 0,65978773 |

|  |  |
| --- | --- |
| Założenie | Wynik |
| Homoskedastyczność | TAK |
| Liniowość | TAK |
| Normalność reszt | TAK |

Model dla powiatów grodzkich spełnia wszystkie założenia Gaussa-Markowa. W tym przypadku parametry modelu będą najefektywniejsze, nieobciążone i zgodne w klasie liniowych estymatorów wektora parametrów modelu.

## Oszacowania parametrów

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *Współczynniki* | *Błąd standardowy* | *t Stat* | *Wartość-p* |
| Przecięcie | -269,7535834 | 85,63133819 | 3,150173629 | 0,002728912 |
| przystanki\_na\_10000 | 1,028479841 | 0,711877285 | 1,444743163 | 0,15464637 |
| cenyMieszkan | 0,001037435 | 0,00010669 | 9,723792267 | 3,32237E-13 |
| zarobki | 0,019729647 | 0,014273134 | 1,382292584 | 0,172905941 |

Interpretacja parametrów modelu:

- zwiększając zmienną ‘przystanki\_na\_10000’ o 1 jednostkę zwiększymy wartość zmiennej objaśnianej o ok. 1,0284 jednostki.

- zwiększając zmienną ‘cenyMieszkan’ o 1 jednostkę zwiększymy wartość zmiennej objaśnianej o ok. 0.001 jednostki.

- zwiększając zmienną ‘zarobki’ o 1 jednostkę zwiększymy wartość zmiennej objaśnianej o ok. 0.019 jednostki.

### Istotność parametrów

* **Istotne predyktory**: Przecięcie i ceny mieszkań są istotnymi statystycznie predyktorami zmiennej zależnej.
* **Nieistotne predyktory**: Liczba przystanków na 10,000 osób i zarobki nie są istotnymi statystycznie predyktorami.

### Ocena modelu dla powiatów grodzkich

Istotność cen mieszkań sugeruje, że są one kluczowym czynnikiem, podczas gdy liczba przystanków i zarobki mają mniejsze znaczenie. Spełnienie założeń Gaussa-Markova potwierdza, że estymatory są wiarygodne i efektywne.

# Model dla powiatów ziemskich

## Poprawność modelu według założeń (Gaussa-Markowa)

|  |  |
| --- | --- |
| Test | p-value |
| Gold.Quand | 0,888616002 |
| Reset-test | 0,34527269 |
| Shapiro-test | 0,002676178 |

|  |  |
| --- | --- |
| Założenie | Wynik |
| Homoskedastyczność | TAK |
| Liniowość | TAK |
| Normalność reszt | NIE |

Model dla powiatów ziemskich nie spełnia wszystkich założenia Gaussa-Markowa. Jednakże:

Założenie o normalności rozkładu składnika losowego nie należy do podstawowych założeń tw. Gaussa-Markowa, a zatem NIE jest niezbędne do uzyskania estymatorów o pożądanych własnościach.[[1]](#footnote-1)

## Oszacowania parametrów

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *Współczynniki* | *Błąd standardowy* | *t Stat* | *Wartość-p* |
| Przecięcie | -35,46324576 | 19,88433137 | 1,783476904 | 0,076032732 |
| przystanki\_na\_10000 | 0,300668044 | 0,062106862 | 4,841140501 | 2,58665E-06 |
| cenyMieszkan | 0,000124282 | 2,33149E-05 | 5,330608544 | 2,6418E-07 |
| zarobki | 0,017552303 | 0,003426779 | 5,122099303 | 7,1153E-07 |

Interpretacja parametrów modelu:

- zwiększając zmienną ‘przystanki\_na\_10000’ o 1 jednostkę zwiększymy wartość zmiennej objaśnianej o ok. 0,30067 jednostki.

- zwiększając zmienną ‘cenyMieszkan’ o 1 jednostkę zwiększymy wartość zmiennej objaśnianej o ok. 0.00012 jednostki.

- zwiększając zmienną ‘zarobki’ o 1 jednostkę zwiększymy wartość zmiennej objaśnianej o ok. 0.01755 jednostki.

### Istotność parametrów

Niespełnienie założenia o normalności rozkładu składnika losowego nie dyskwalifikuje wyników testów. Statystyka t-studenta [….] ma w dużych próbach rozkład asymptotycznie normalny […].[[2]](#footnote-2)

Dla powiatów ziemskich dysponujemy dużą ilością danych (203), dlatego przyjmujemy poprawne wyniki testów istotności.

* **Istotne predyktory**: Ceny mieszkań, liczba przystanków na 10,000 osób oraz zarobki są istotnymi statystycznie predyktorami zmiennej zależnej.
* **Nieistotne predyktory**: Przecięcie nie jest istotnym statystycznie predyktorem.

### Ocena modelu dla powiatów ziemskich

Wszystkie parametry poza przecięciem wydają się być istotne. Bazując na podanych przez nas źródłach, model jest w stanie pozostać wiarygodny i efektywny.

# Model dla wszystkich powiatów

## Poprawność modelu według założeń (Gaussa-Markova)

|  |  |
| --- | --- |
| Test | p-value |
| Gold.Quand | 0 |
| Reset-test | 0,07867888 |
| Shapiro-test | 0 |

|  |  |
| --- | --- |
| Założenie | Wynik |
| Homoskedastyczność | NIE |
| Liniowość | TAK |
| Normalność reszt | NIE |

Model dla powiatów ziemskich nie spełnia wszystkich założenia Gaussa-Markova.

Biorąc pod uwagę nasze wcześniejsze odwołania do książki, przyjmijmy, że defektem tego modelu jest wyłącznie heteroskedastyczność. Jej skutkiem jest, to, że estymatory nie są już najefektywniejsze (w klasie estymatorów liniowych i nieobciążonych).

## Oszacowania parametrów

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *Współczynniki* | *Błąd standardowy* | *t Stat* | *Wartość-p* |
| Przecięcie | 138,3219876 | 68,23895861 | 2,027023717 | 0,043400627 |
| gestosc | 0,040274176 | 0,013104761 | 3,07324758 | 0,002279776 |
| odsetek65 | -6,154682786 | 3,039116906 | 2,025154997 | 0,043593104 |
| cenyMieszkan | 0,000530761 | 9,17929E-05 | 5,782154889 | 1,60758E-08 |

Interpretacja parametrów modelu:

- zwiększając zmienną ‘gestosc’ o 1 jednostkę zwiększymy wartość zmiennej objaśnianej o ok. 0,30067 jednostki.

- zwiększając zmienną ‘odsetek65’ o 1 jednostkę zmniejszymy wartość zmiennej objaśnianej o ok. -6,15 jednostek.

- zwiększając zmienną ‘cenyMieszkan o 1 jednostkę zwiększymy wartość zmiennej objaśnianej o ok. 0.0005 jednostki.

### Istotność parametrów

W modelu dla wszystkich powiatów wszystkie zmienne są istotne.

### Ocena modelu dla powiatów grodzkich

Bazując na podanych przez nas źródłach, model jest w stanie pozostać wiarygodny i efektywny.

# Test Chowa

**Hipotezy:**

**Statystyka:**

|  |  |
| --- | --- |
| F\_obl | F\_kryt |
| 2,1659 | 2,3972 |

**Rezultat:**

=> Przyjmujemy hipotezę zerową, mówiącą, że różnice między grupami nie miały istotnego wpływu na strukturę zależności między zmiennymi. **Oznacza to, że wspólny model jest stabilny.**

1. M. Gruszczyński, T.Kuszewski, M. Podgórska, *Ekonometria i badania operacyjne*, Warszawa 2009, s. 88. [↑](#footnote-ref-1)
2. M. Gruszczyński, T.Kuszewski, M. Podgórska, *Ekonometria i badania operacyjne*, Warszawa 2009, s. 89. [↑](#footnote-ref-2)