Ekonometria

Robert Pietrzykowski
email: robert_pietrzykowski@sggw.pl
www.ekonometria.info

Na dziś…

- Sprawy bieżące

Praca własna…

- Założenia klasycznego modelu regresji liniowej (KMRL) [Górecki 2010 s. 25-31]
- METODA NAJMNIEJSZYCH KWADRATÓW [Górecki 2010 s.33 – 37]

Wybór modelu

Oszacowanie parametrów modelu

Zbadanie istnienia zależności

Ocena jakości dopasowania modelu

Sprawdzenie poprawności modelu

Prognoza, predykcja i etc.

Zapis macierzowy

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

gdzie:

$$\boldsymbol{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \quad \boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{p1} \\ 1 & x_{12} & \cdots & x_{p2} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_{1n} & \cdots & x_{pn} \end{bmatrix} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

oraz $\boldsymbol{\beta}^T = [\, \beta_0 \quad \beta_1 \quad \cdots \quad \beta_p \,]$

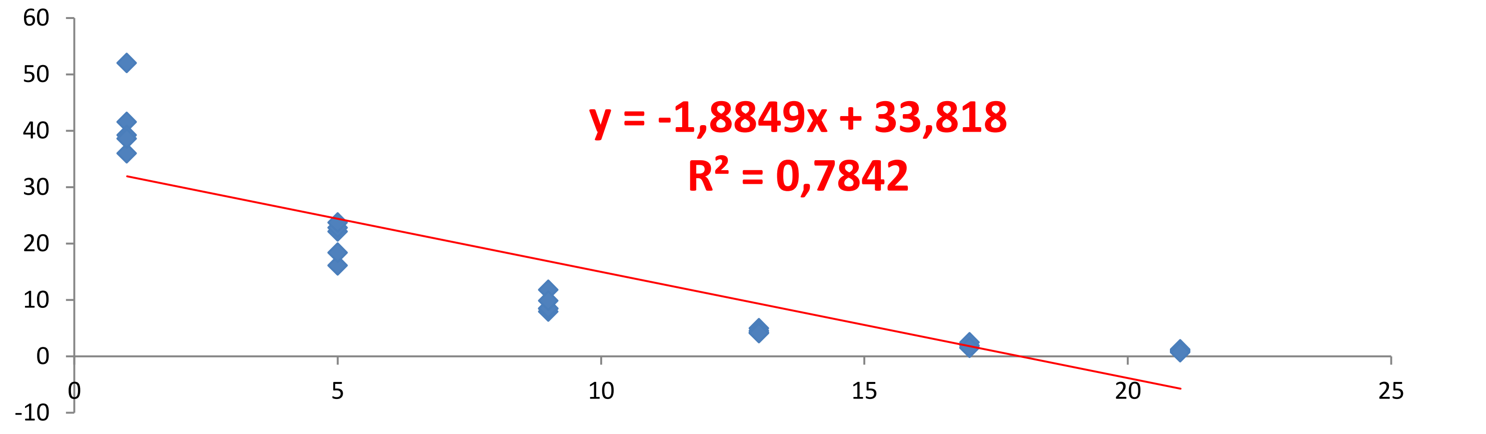
$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{21} & x_{31} & \cdots & x_{K1} \\ 1 & x_{22} & x_{32} & \cdots & x_{K2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{2n} & x_{3n} & \cdots & x_{Kn} \end{bmatrix} \times \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

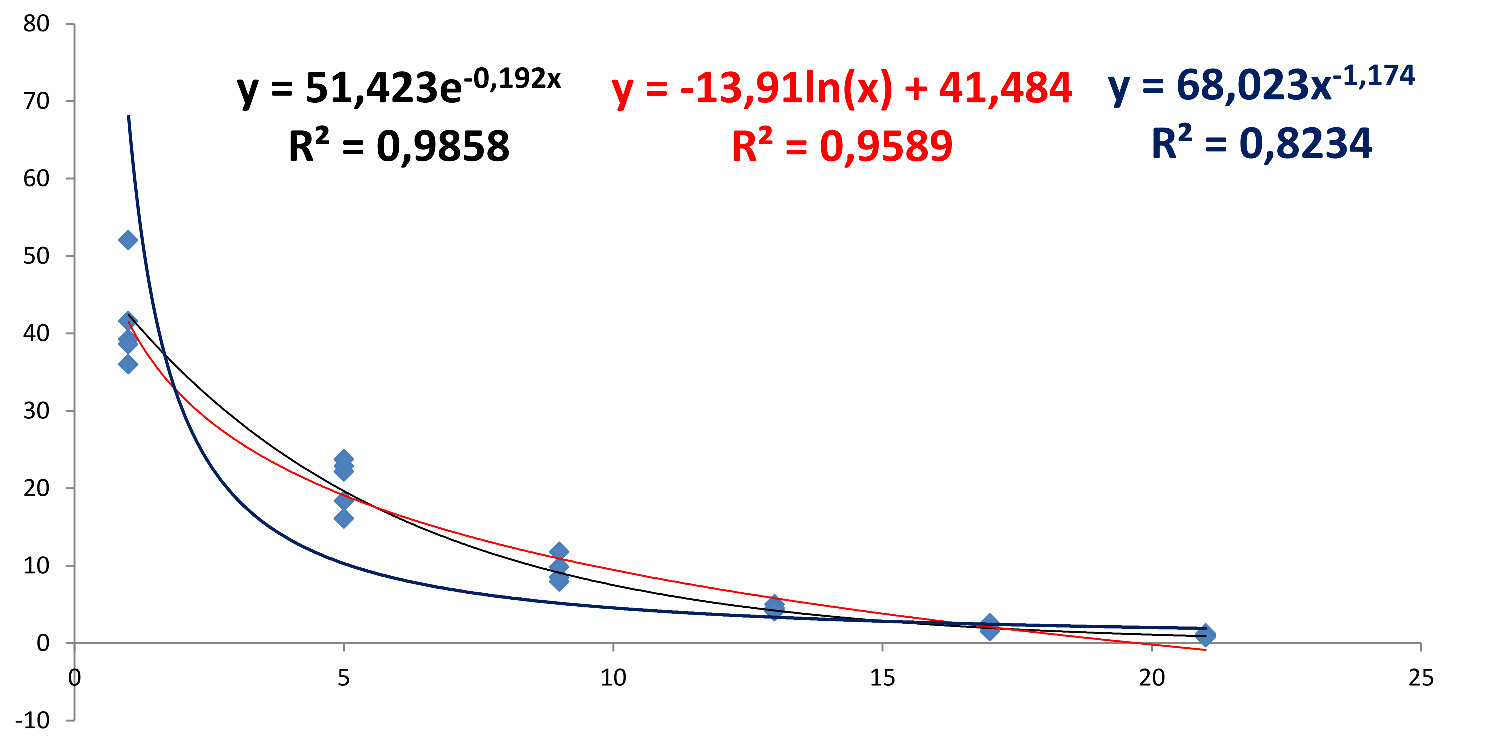
$$
\begin{bmatrix}
n & \sum x_1 & \cdots & \sum x_K \\
\sum x_1 & \sum x_1^2 & \cdots & \sum x_1 x_K \\
\vdots & \vdots & & \vdots \\
\sum x_K & \sum x_K x_1 & \cdots & \sum x_K^2
\end{bmatrix}
\times
\begin{bmatrix}
b_1 \\ b_2 \\ \vdots \\ b_K
\end{bmatrix}
=
\begin{bmatrix}
\sum_{i=1}^{n} y_i \\
\sum_{i=1}^{n} x_{2i} y_i \\
\\
\sum_{i=1}^{n} x_{Ki} y_i
\end{bmatrix}
$$

Rozważamy zależność między ceną a popytem na pewien towar.

1. DOBRAĆ MODEL
2. ZBADAĆ ZALEŻNOŚĆ
3. PRZEPROWADZIĆ ANALIZĘ RESZTOWĄ
4. O ILE ŚREDNIO ZMIENI SIĘ POPYT JEŻELI CENĘ OBNIŻYMY CENĘ O 1 ZŁ.
5. ODPOWIEDZIEĆ NA PYTANIE JAKIEGO ŚREDNIEGO POPYTU MOŻNA SIĘ PODZIEWAĆ PRZY CENIE RÓWNEJ 10
6. ODPOWIEDZIEĆ NA PYTANIE PRZY JAKIEJ CENIE OSIAGANY JEST MAKSYMALNY PRZYCHÓD

Cena	Popyt
1	35,969
1	41,55
1	39,213
1	52,014
1	38,579
5	16,068
5	22,141
5	22,821
5	23,705
5	18,344
9	7,918
9	8,439
9	8,472
9	11,78
9	9,832
13	4,092
13	4,992
13	4,221
13	4,503
13	4,102
17	1,904
17	1,523
17	1,523
17	2,489
17	1,545
21	0,95
21	0,849
21	1,078
21	1,209
21	0,722





Regresja prosta

Funkcja regresji zależna tylko od jednego argumentu

Nazwa funkcji	Wzór funkcji	Model
Liniowa	$a + bx$	$y = a + bx$
Potęgowa	ax^b	$\ln y = \ln a + b \ln x$
Wykładnicza	$\exp(a + bx)$	$\ln y = a + bx$
Typu S	$\exp(a + \frac{b}{x})$	$\ln y = a + b\frac{1}{x}$
Hiperboliczna	$\frac{1}{a+bx}$	$\frac{1}{y} = a + bx$
Podwójnie hiperboliczna	$\frac{1}{a+b/x}$	$\frac{1}{y} = a + b\frac{1}{x}$
Pierwiastkowa	$a + b\sqrt{x}$	$y = a + b\sqrt{x}$
Logarytmiczna	$a + b \ln x$	$y = a + b \ln x$

Wyniki dopasowania modeli.

f. regresji	$\beta_1 x + \beta_0$	$\beta_1/x + \beta_0$	$\beta_0 x^{\beta_1}$	$\beta_1 ln(x) + \beta_0$	$\exp(\beta_1 x + \beta_0)$
$\hat{\beta}_1$	−1.885	39.832	−1.174	−13.911	−0.192
$\hat{\beta}_0$	33.819	3.164	4.220	41.484	3.940
S_{β_1}	0.187	2.971	0.103	0.544	0.004
S^2	48.871	30.520	0.332	9.312	0.027
F_{emp}	101.774	179.804	130.546	653.077	1939.484
D	0.784	0.865	0.823	0.959	0.986
RSS	1368.390	854.563	9.293	260.736	0.749
$VarB.cz.$	208.666	208.666	0.667	208.666	0.667
$VarB.d.$	1159.724	645.897	8.626	52.070	0.082

Potęgowa	ax^b	$\ln y = \ln a + b \ln x$
Wykładnicza	$\exp(a + bx)$	$\ln y = a + bx$
Logarytmiczna	$a + b \ln x$	$y = a + b \ln x$

Analiza resztowa

1) Badanie normalności rozkładu reszt:

2) Badanie homoskedastyczności wariancji:

3) Badanie autokorelacji reszt:

4) Sprawdzanie specyfikacji modelu

$$LN(POPYTU) = 51{,}423\, e^{-0{,}192x}$$

$$LN(POPYTU) = e^{3{,}940 - 0{,}192x}$$

O ILE ŚREDNIO ZMIENI SIĘ POPYT JEŻELI CENĘ OBNIŻYMY CENĘ O 1 ZŁ.

Przedział ufności dla współczynnika regresji:

$$\left(\widehat{\beta}_j - t^{\alpha}_{n-p-1} S_{\beta_j},\ \widehat{\beta}_j + t^{\alpha}_{n-p-1} S_{\beta_j}\right)$$

$$\beta_1 \in (-0{,}20145; -0.18354)$$

Jeżeli obniżymy cenę o 1 zł to możemy się spodziewać wzrostu popytu średnio w zakresie od 18 do 20 groszy.

ODPOWIEDZIEĆ NA PYTANIE JAKIEGO ŚREDNIEGO POPYTU MOŻNA SIĘ PODZIEWAĆ PRZY CENIE RÓWNEJ 10

OBSZAR UFNOŚCI

$$E(Y|x) \in (\widehat{y}(x) - t(\alpha; n-2)S_Y ; \widehat{y}(x) + t(\alpha; n-2)S_Y)$$

$$S_Y^2 = S^2 \left(\frac{1}{n} + \frac{(x - \bar{X})^2}{\mathrm{var}X} \right)$$

OBSZAR PREDYKCJI

$$Y(x) \in (\widehat{y}(x) - t(\alpha; n-2)S_{y(x)} ; \widehat{y}(x) + t(\alpha; n-2)S_{y(x)})$$

$$S_{y(x)}^2 = S^2 \left(1 + \frac{1}{n} + \frac{(x - \bar{X})^2}{\mathrm{var}X} \right)$$

Aby oszacować wartość funkcji regresji w punkcie

$$\boldsymbol{x}'_\star = (1, x_{\star 1}, \cdots, x_{\star p})$$

przyjmujemy $\boldsymbol{\lambda}' = \boldsymbol{x}'_\star$

ocena punktowa:

$$\widehat{\bar{Y}_{(x_\star)}} = \boldsymbol{x}'_\star \widehat{\boldsymbol{\beta}}$$

OBSZAR UFNOŚCI

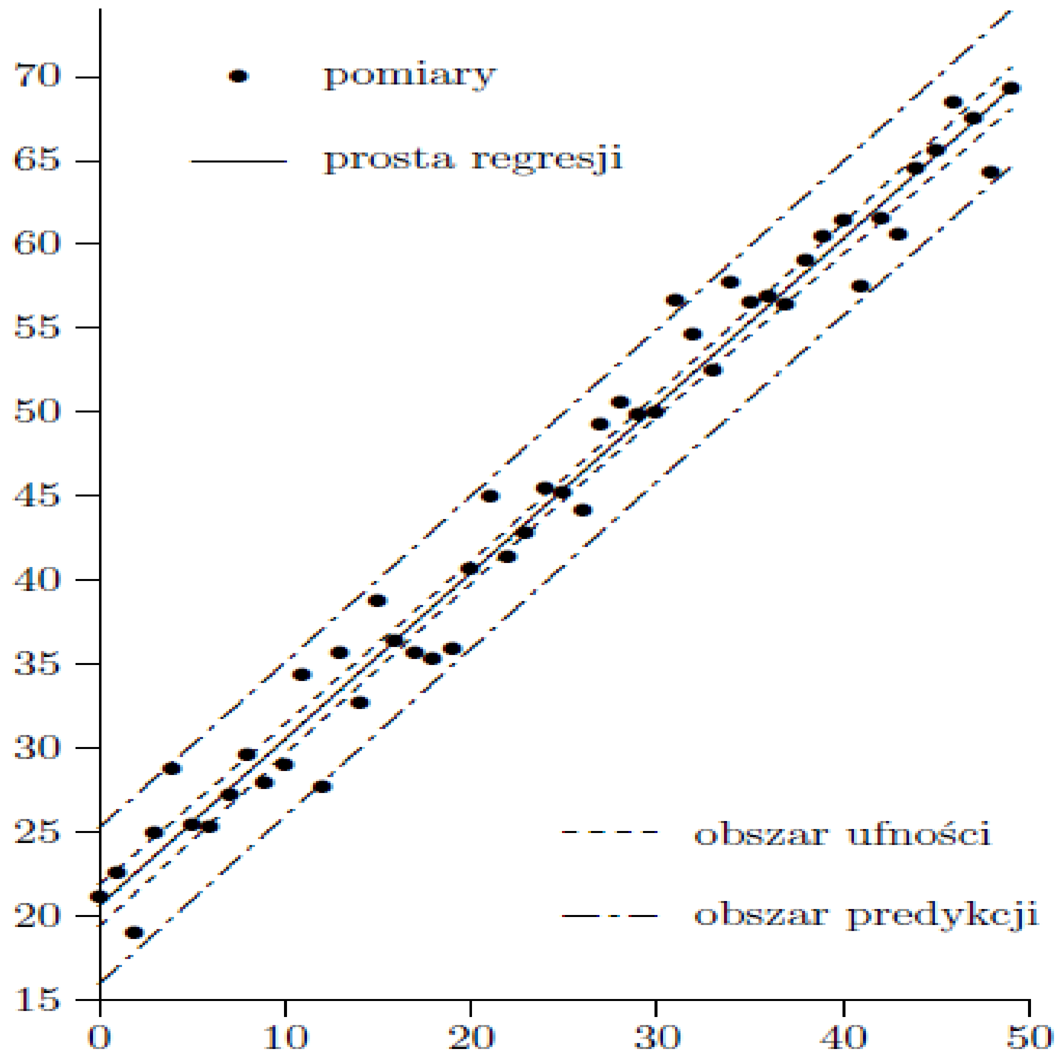
$$\left(\boldsymbol{x}_\star'\widehat{\boldsymbol{\beta}} - t_{n-p-1}^{\alpha} S_{\bar{\mathrm{Y}}_{(x_\star)}},\ \boldsymbol{x}_\star'\widehat{\boldsymbol{\beta}} + t_{n-p-1}^{\alpha} S_{\bar{\mathrm{Y}}_{(x_\star)}}\right)$$

$$S_{\bar{\mathrm{Y}}_{(x_\star)}}^2 = s^2 \boldsymbol{x}_\star'(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{x}_\star$$

OBSZAR PREDYKCJI

$$\left(\boldsymbol{x}_\star'\widehat{\boldsymbol{\beta}} - t_{n-p-1}^{\alpha} S_{\mathrm{Y}_{(x_\star)}},\ \boldsymbol{x}_\star'\widehat{\boldsymbol{\beta}} + t_{n-p-1}^{\alpha} S_{\mathrm{Y}_{(x_\star)}}\right)$$

$$S_{\mathrm{Y}_{(x_\star)}}^2 = s^2 \left(\boldsymbol{x}_\star'(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{x}_\star + 1\right)$$


pomiary
prosta regresji
obszar ufności
obszar predykcji
70
65
60
55
50
45
40
35
30
25
20
15
0
10
20
30
40
50

ODPOWIEDZIEĆ NA PYTANIE JAKIEGO ŚREDNIEGO POPYTU MOŻNA SIĘ PODZIEWAĆ PRZY CENIE RÓWNEJ 10

OBSZAR UFNOŚCI

$$E(Y|x) \in (\widehat{y}(x) - t(\alpha; n-2)S_Y; \widehat{y}(x) + t(\alpha; n-2)S_Y)$$

$$S_Y^2 = S^2 \left(\frac{1}{n} + \frac{(x - \bar{X})^2}{\mathrm{var} X} \right)$$

E(Y|X = 10) = 7,5019, a przedział (7,4414; 7,5624)

Jeżeli cenę ustalimy na 10zł to możemy spodziewać się średniego popytu w wysokości od 7,4414 do 7,5624

PRZY JAKIEJ CENIE OSIAGANY JEST MAKSYMALNY PRZYCHÓD

- Funkcja przychodu

$$f(x) = x \cdot e^{3.9401-0.1925x}$$

$$f(x)' = x' \cdot e^{3.9401-0.1925x} + x \cdot [e^{3.9401-0.1925x}]'$$

$$f(x)' = 1 \cdot e^{3.9401-0.1925x} + (-0.1925x) \cdot e^{3.9401-0.1925x} = 0$$

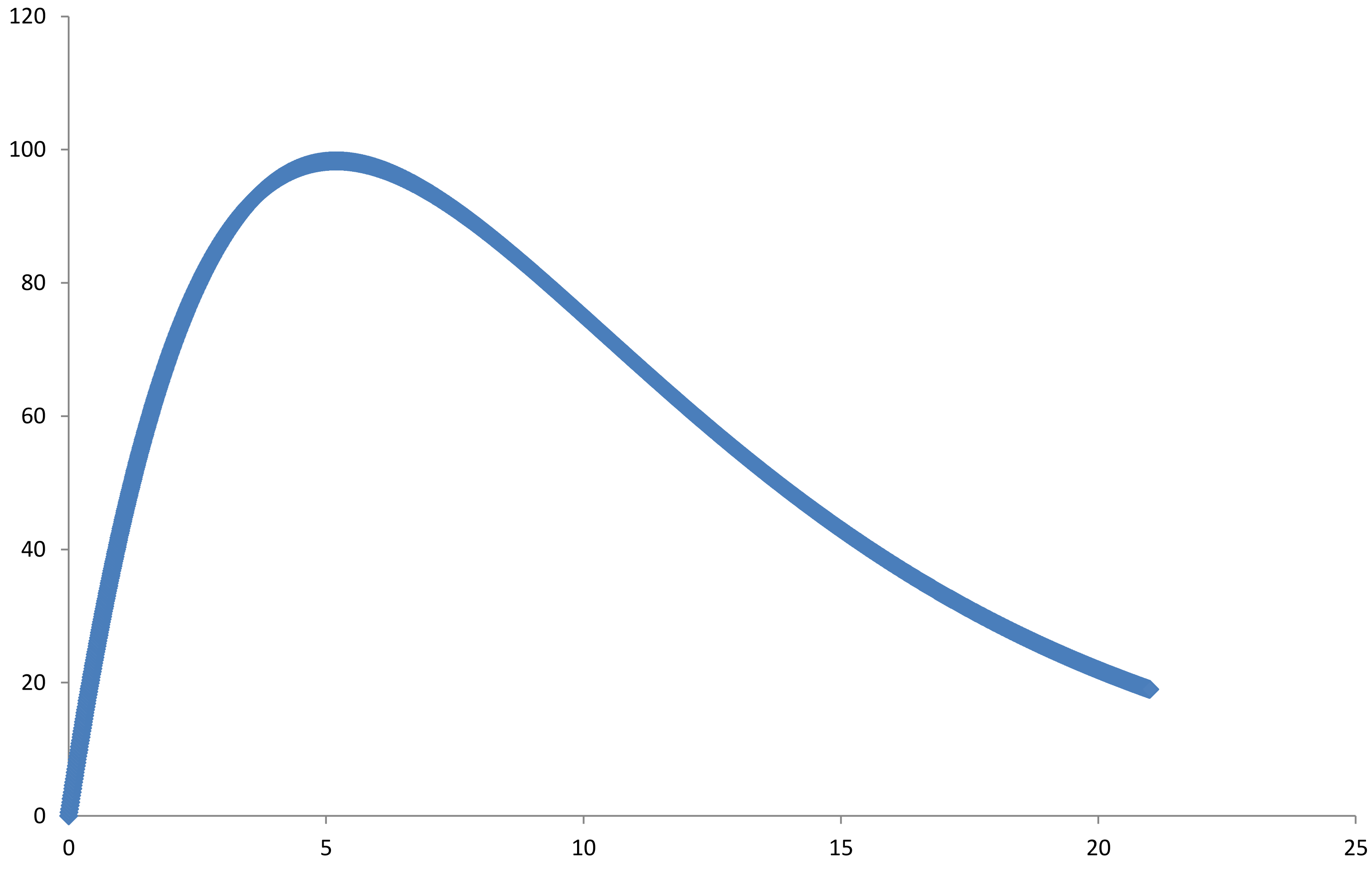
$$-0.1925x \cdot e^{3.9401-0.1925x} = -1 \cdot e^{3.9401-0.1925x}$$

$$-0.1925x = -1$$

$$x = \frac{-1}{-0.1925}$$

$$x = 5.2083$$

$$f(x) = x \cdot e^{3.9401 - 0.1925x}$$



PYTANIA

1. Podaj postać liniową funkcji produkcji Cobba-Douglasa.
2. Wyjaśnij pojęcie modele linearyzowalne.
3. Jak jest różnica miedzy obszarem ufności i obszarem predykcji.
4. Zapisz postać obszaru ufności dla x=10 wiedzą, że y=12 – 0,2x
5. Zinterpretuj następujący obszar predykcji (12,02; 17,12) dla zależności pomiędzy zużyciem paliwa samochodu, a masą ładunku wynoszącą 6 ton.
6. Dlaczego obszar ufności jest zawsze węższy od obszaru predykcji.
7. Narysuj schematycznie na wykresie funkcję regresji oraz granice obszaru ufności i predykcji.
8. Czy interpretacja współczynnika regresji jest poprawna dla każdej funkcji regresji zależnej od jednego argumentu. Odpowiedź uzasadnij.
9. Narysuj wykres P-P, na podstawie którego można stwierdzić, że badana zmienna losowa nie ma rozkładu normalnego.
10. Badając zależność pomiędzy popytem na pewien produkt, a wielkością środków przeznaczanych na reklamę TV. Uzyskano przedział ufności dla współczynnika regresji (-1,21; 2,32). Co możesz powiedzieć o badanej zależności. Zinterpretuj współczynnik regresji.