ZADANIA – Zestaw 5

Zadanie 5.1

X – cena (zł), *Y* – popyt (tys. szt.). Mając dane

x	1,5	2	2,5	3	3,5	4	4,5	5
y	44	43	43	37	36	34	35	35

Oblicz współczynnik korelacji Pearsona.

Oblicz współczynnik korelacji Spearmana.

Wyznaczyć równanie prostej regresji.

Oceń jej dopasowanie do danych statystycznych.

Jaki będzie popyt przy cenie 6zł.?

(odp. r = -0,91; Q = -0,89; $\hat{y} = 48{,}72 - 3{,}12x$, $R^2 = 0{,}82$)

Zadanie 5.2

Y – plony pszenicy z ha (w q),
X – zużycie nawozów mineralnych na 1 ha (w kg czystego składnika NPK).
Dane z lat 1990 ÷ 1997:

Rok	1990	1991	1992	1993	1994	1995	1996	1997
X	150	160	170	180	190	195	200	220
Y	24	30	32	30	32	30	35	40

Obliczyć współczynnik korelacji Pearsona. Wyznaczyć równanie prostej regresji.
Obliczyć współczynnik determinacji i podać jego interpretację.
Wyznaczyć prognozę wielkości plonów pszenicy na rok 1998 jeśli planowane zużycie nawozów w tym roku ma wynosić 210 kg na ha.

(odp. $r = 0{,}87$, $\hat{y} = -0{,}481 + 0{,}175x$)

Zadanie 5.3

Producent napojów chłodzących zgromadził dane o wielkości zamówień hurtowni – *Y* (w tys. l.) i średniej temperaturze dobowej w okresie lipiec– sierpień dla przypadkowo wybranych 10 dni – *X* (w stopniach C):

Średnia temp. dobowa	18	24	29	20	35	18	14	27	30	22
Wielkość zamówień	50	93	119	60	160	52	35	105	120	71

Obliczyć współczynnik korelacji Pearsona. Obliczyć współczynnik korelacji rangowej. Czy zależność między wielkością zamówień napojów chłodzących, a temperaturą dobową jest silna? Zbudować model regresji. Przedstawić graficznie badany związek.
(odp. $r = 0{,}99$, $\hat{y} = -56{,}5 + 6{,}034x$)

Zadanie 5.4

Na podstawie danych:

X	1,5	2,4	3,9	5,6	6,5	8,0	8,3	9,4	9,5	10,3	11,0
Y	2,1	1,2	4,2	3,0	6,1	9,9	13,1	14,2	17,0	19,2	25,0

wyznaczono równanie regresji

$$\hat{Y} = -5{,}17 + 2{,}25X$$

Narysować wyznaczoną prostą regresji i zaznaczyć powyższe dane statystyczne.

Oblicz średnie błędy współczynników tej prostej. Który współczynnik jest obliczony z wiekszą dokładnością?
Czy cechy *X* i *Y* są skorelowane dodatnio?
Oblicz współczynnik korelacji Pearsona,
Oblicz współczynnik korelacji Spearmana.

(odp. $r = 0{,}924$; $Q = 0{,}98$)

Zadanie 5.5

X – miejsce w konkursie, *Y* – liczba udzielonych wywiadów. Mając dane

uczestnik	A	B	C	D	E	F	G	H
x	1	2	3-4	3-4	5-6-7	5-6-7	5-6-7	8
y	10	10	6	8	5	6	5	3

Oblicz współczynnik korelacji Spearmana.

Zadanie 5.6

Korzystając z danych zgrupowanych w tablicy korelacyjnej wyznacz współczynnik korelacji i równanie prostej regresji cechy *Y* względem *X*.

X \ *Y*	5	6	7	8	$n_{i\cdot}$
5	0	2	0	8	
6	0	4	2	9	
7	2	12	3	1	
8	21	14	0	0	
9	1	0	0	0	
$n_{\cdot j}$					

(odp. $r = -0{,}79$, $\hat{y} = 11{,}86 - 0{,}8x$)

Zadanie 5.7

Korzystając z danych zgrupowanych w tablicy korelacyjnej wyznacz współczynnik korelacji i równanie prostej regresji cechy *Y* względem *X*.
Wyznacz empiryczne linie regresji (regresja I rodzaju) i stosunki korelacyjne. Oceń graficznie i liczbowo krzywoliniowość tej regresji.

X \ *Y*	1	2	3	4	$n_{i\cdot}$
2	5	0	0	0	
4	0	5	3	0	
6	0	3	5	0	
8	0	0	5	4	
$n_{\cdot j}$					

(odp. $r = 0{,}84$, $\hat{y} = 0{,}6 + 0{,}36x$)

Zadanie 5.8

Y – wydatki miesięczne na żywność w rodzinie (setki. zł),
X – dochody miesięczne na jednego członka rodziny (tys. zł),
Zbadano $n = 9$ rodzin, otrzymano następujące wyniki:

$$\sum_{i=1}^{9}(x_i - \bar{x})(y_i - \bar{y}) = 12{,}02 \;;\quad \sum_{i=1}^{9} x_i = 6{,}4 \;;\quad \sum_{i=1}^{9} y_i = 124 \;;$$

$$\sum_{i=1}^{9} x_i^2 = 5{,}52 \;;\qquad \sum_{i=1}^{9}(y_i - \bar{y})^2 = 151{,}56$$

Obliczyć współczynnik korelacji Pearsona. Wyznaczyć równanie prostej regresji. Obliczyć R^2, $s_e^2, s(b_i) \quad i = 0,1$.

(odp. $R^2 = 0,98$; $\hat{y} = 4,95 + 12,41x$; $s_e = 0,58$, $s(b_0) = 0,46$; $s(b_1) = 0,7$)

Zadanie 5.9

Na podstawie danych:

X	2,7	4,6	6,3	7,8	9,2	10,6	12	13,4	14,7
Y	17,0	16,2	13,3	13,0	9,7	9,9	6,2	5,8	5,7

otrzymano następujące wyniki:

$$\sum_i x_i y_i = 735,7 \quad ; \qquad \sum_i x_i = 81,3 \quad ; \qquad \sum_i y_i = 96,8 \quad ;$$

$$\sum_i x_i^2 = 865,63 ; \qquad \sum_i y_i^2 = 1194$$

Obliczyć współczynnik korelacji Pearsona. Wyznaczyć równanie prostej regresji.
Wyznaczyć prognozę punktową dla $x = 15$ i ocenić jej dokładność.

(odp. $r = -0,979$, $\hat{y} = 20,3 - 1,06x$)

Zadanie 5.10

Y – wydajność (szt./h), X – czas od zainstalowania maszyny (miesiące),
Zbadano $n = 11$ maszyn, otrzymano następujące wyniki:

$$\sum_i x_i y_i = 682,25 \quad ; \qquad \sum_i x_i = 86 \quad ; \quad \sum_i y_i = 98,26 \quad ;$$

$$\sum_i x_i^2 = 868 ; \quad \sum_i y_i^2 = 1087,91$$

Obliczyć współczynnik korelacji Pearsona. Wyznaczyć równanie prostej regresji. Obliczyć R^2, $s_e^2, s(b_i) \quad i = 0,1$.

(odp. $r = -0,434$, $\hat{y} = 12,37 - 0,44x$)

Zadanie 5.11

W modelu liniowej regresji zadłużenia (Y – tys. PLN) podmiotów gospodarczych z tytułu kredytów względem wartości produkcji sprzedanej (X – mln PLN) dla 80 podmiotów z sektora produkcyjnego otrzymano:

$\text{cov}(X, Y) = -2,88$; $\quad \bar{x} = 5,5$; $\quad s(X) = 0,6$; $\quad \bar{y} = 183,0$; $\quad s(Y) = 8,0$

Wyznaczyć prognozę punktową dla $x = 9$ i ocenić jej dokładność.

(odp. $r = -0,6$, $\hat{y} = 227 - 8x$, $\hat{Y}(9) = 155$; $S_\tau = 7,77$)

Zadanie 5.12

Badając współzależność między zmiennymi X i Y, uzyskano w próbie losowej o liczebności $n = 20$ następujące rezultaty:

$\bar{x} = 5$; $\qquad \sum x_i^2 = 520$;

$s(Y) = 3$; $\qquad v(Y) = 20\%$;

$\sum x_i y_j = 1450$

Obliczyć współczynnik korelacji liniowej Pearsona, wyznaczyć prostą regresji.

(odp. $r = -0{,}83$, $\hat{y} = 27{,}5 - 2{,}5x$)

Zadanie 5.13

X – liczba punktów uzyskanych przez studentów ze statystyki w połowie semestru (pierwsze kolokwium),
Y – liczba punktów uzyskanych przez studentów ze statystyki na koniec semestru (drugie kolokwium),

X	71	49	80	73	93	85	58	82	64	32	87	80
Y	83	62	76	77	89	74	48	78	76	51	73	89

Obliczyć współczynnik korelacji liniowej Pearsona.
Czy cechy X i Y są skorelowane dodatnio?
Wyznaczyć prostą regresji.
Ile średnio punktów powinien otrzymać student z drugiego kolokwium, jeśli pierwsze kolokwium napisał na 84 punkty?

(odp. cechy skorelowane dodatnio, $\hat{y} = 31{,}6 + 0{,}58x$, około 80 punktów)

Zadanie 5.14

Wiedząc, że wariancja resztowa modelu regresji liniowej obliczona dla 20 danych wynosi 5, a wariancja zmiennej Y wynosi 25 wyznacz współczynnik determinacji tego modelu.

(odp. 0,82)

Zadanie 5.15

Udowodnij równość występującą we wzorze na standardowy błąd prognozy

$$s_\tau = s_e \sqrt{1 + \frac{1}{n} + \frac{(x_\tau - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}} = s_e \sqrt{1 + \frac{\sum_{i=1}^{n} x_i{}^2 + n x_\tau^2 - 2x_\tau \sum_{i=1}^{n} x_i}{n\sum_{i=1}^{n} x_i{}^2 - \left(\sum_{i=1}^{n} x_i\right)^2}}$$

Zadanie 5.16

Y – wydatki miesięczne na żywność w rodzinie (tys. zł),
X – dochody miesięczne na jednego członka rodziny (tys. zł),
Zbadano n = 200 rodzin, otrzymano następujące wyniki:

$$\sum_{i=1}^{200}(x_i - \bar{x})(y_i - \bar{y}) = 66{,}5; \qquad \sum_{i=1}^{200} x_i = 300; \qquad \sum_{i=1}^{200} y_i = 120;$$

$$\sum_{i=1}^{200} x_i^2 = 5000; \qquad \sum_{i=1}^{200} y_i^2 = 760$$

Obliczyć współczynnik korelacji. Wyznaczyć prostą regresji z próby. Ocenić dopasowanie prostej regresji do danych empirycznych.

(odp. $r = 0{,}037$; $\hat{Y} = 0{,}58 + 0{,}015X$, $s_e = 1{,}86$).

Zadanie 5.17
Badano dwie cechy.
Pobrano próbę 10 elementową i otrzymano następujące wyniki:

$$\sum_{i=1}^{10} x_i y_i = 792{,}92\,; \qquad \sum_{i=1}^{10} x_i = 86{,}7\,; \quad \sum_{i=1}^{10} y_i = 88{,}8\,;$$

$$\sum_{i=1}^{10} x_i^2 = 771{,}35\,; \qquad \sum_{i=1}^{10} y_i^2 = 819{,}34$$

Obliczyć współczynnik korelacji. Wyznaczyć prostą regresji z próby. Ocenić dopasowanie prostej regresji do danych empirycznych.
Wyznaczyć prognozę punktową dla $x = 10$ i ocenić jej dokładność.

(odp. $r = 0{,}936$).

Zadanie 5.18
Badano zależność między liczbą godzin poświęconych przygotowaniu do egzaminu (cecha X) a liczbą punktów uzyskanych na egzaminie (cecha Y).
Dla próby 10 elementowej otrzymano następujące wyniki:

X	4	9	10	14	4	7	12	22	1	17
Y	31	58	65	73	37	44	60	91	21	84

Obliczyć współczynnik korelacji. Wyznaczyć prostą regresji z próby. Ocenić dopasowanie prostej regresji do danych empirycznych.

(odp. $y = 3{,}47x + 21{,}69$).

Zadanie 5.19
Badano zależność między wielkością produkcji a kosztami całkowitymi produkcji pewnego wyrobu.
Dla próby 10 elementowej otrzymano następujące wyniki (X – wielkość produkcji w setkach sztuk, Y – koszty w tysiącach złotych):

X	1,0	1,5	2,0	3,0	4,0	10,0	16,0	20,0	25,0	32,0
Y	1,0	2,5	3,0	3,5	4,0	5,0	6,0	7,0	8,0	9,0

Obliczyć współczynnik korelacji. Wyznaczyć prostą regresji z próby. Ocenić dopasowanie prostej regresji do danych empirycznych.

(odp. $y = 0{,}29X + 1{,}9$).

Zadanie 5.20
Y – wartość sprzedaży (w mln złotych),
X – wydatki na reklamę (w dziesiątki tys. zł).

Y	11	13	16	12	13	19	22	14	15	25	28	16	17	31	34	18
X	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8	8

Obliczyć współczynnik korelacji. Wyznaczyć prostą regresji z próby. Ocenić dopasowanie prostej regresji do danych empirycznych.
Wyznaczyć prognozę punktową dla $x = 9$ i ocenić jej dokładność.

(odp. $r = 0{,}67$; prosta regresji $Y = 10 + 2X$).

Zadanie 5.21

W modelu $\hat{Y} = 2 - 4X$, średnia zmiennej X wynosi 5.

Oblicz średnią zmiennej Y.

Zadanie 5.22

W modelu $\hat{Y} = 2 - 4X$, średnia zmiennej X wynosi 4, odchylenie standardowe zmiennej Y wynosi 5, współczynnik zmienności zmiennej X wynosi 0,3.

Oblicz współczynnik korelacji.

(odp. r = 0,96).

Zadanie 5.23

Wiedząc, że S(Y) = 5,16 i reszty malejącej prostej regresji są równe:

-2,164; -0,461; 0,242; 0,945; 0,648; 3,352; -0,945; 0,758; 0,461; -2,836

Oblicz współczynnik korelacji. Oblicz odchylenie resztowe.

(odp. R^2 = 0,898; S_e = 1,84).

Zadanie 5.24

Y - udział braków w produkcji pewnej firmy (w promilach),

Dane z lat 1992-98:

Rok	**1992**	**1993**	**1994**	**1995**	**1996**	**1997**	**1998**
Y	13	12	10	9	8	6	5

Obliczyć współczynnik korelacji. Wyznaczyć prostą regresji z próby. Ocenić dopasowanie prostej regresji do danych empirycznych.
Wyznaczyć prognozę punktową udziału braków w produkcji w roku 2000. Oceń błąd tej prognozy.

(odp. R^2 = 0,99, $\hat{y} = 14{,}43 - 1{,}36t$)

Zadanie 5.25

Funkcja trendu wyznaczona na podstawie danych o wielkości sprzedaży (w tonach) w pewnej firmie w ostatnich 9 miesiącach 2009 r. ma postać:

$\hat{y}_t = 150 + 7{,}5t$ S_e = 7,2 tony, R^2 = 0,95

a) wyznaczyć prognozę na styczeń 2010,
b) ocenić dopuszczalność zbudowanej prognozy,

L.Kowalski, 23.04.2010