

## Analiza korelacji i regresji

**8.1.** Badano zależność między ilością godzin przebywania samolotu w powietrzu („nalot” lotniczy) a ilością wypadków. Na podstawie zebranych danych z pewnego okresu czasu zbadać czy taka zależność istnieje, a jeżeli istnieje, to opisać ją za pomocą liniowej funkcji regresji.

„nalot”	167	167	155	118	159	141	130	131	141	135	145	147	128	137	138
wypadki	135	129	118	134	120	109	@92	129	138	132	138	135	111	123	@99

**8.2.** Zbadano w dziesięciu wylosowanych zakładach przemysłowych wielkość zaplanowanego i wykonanego funduszu na akcję socjalną. Zbadać, czy istnieje zależność między badanymi cechami. Jeżeli taka zależność istnieje, to opisać ją za pomocą liniowej funkcji regresji.

planowany	3.60	4.65	5.20	1.86	3.06	1.36	2.46	3.93	5.80	6.35
wykonany	3.56	4.59	5.13	1.84	3.02	1.35	2.43	3.89	5.72	6.26

**8.3.** Istnieje podejrzenie że ludzie dzielą się na humanistów i matematyków tzn. jeśli ktoś jest dobry z przedmiotów humanistycznych to z matematyką może już mieć problemy. Wylosowano ośmiu uczniów z czwartej klasy liceum i obliczono dla nich średnie z ocen semestralnych z języka polskiego i z matematyki:

J. polski	3.4	2.9	4.3	3.8	3.3	4.7	3.6	3.9
Matematyka	4.6	4.4	3.6	3.2	3.8	3.5	4.6	3.1

Interesuje nas, czy oceny uczniów potwierdzają wspomniane twierdzenie. Jeśli tak to proszę wyznaczyć równanie regresji. Jaki jest przewidywana średnia ocena z matematyki dla ucznia którego średnia z języka polskiego wynosi 5.0?

**8.4.** Poniższe dane z dziesięciu poletek dotyczą efektywności nawożenia łąk azotem (w kg siana na 1 kg N) w zależności od poziomu nawożenia azotem:

$$\bar{x} = 40, \quad \bar{y} = 16, \quad \sum x_i^2 = 20200, \quad \sum y_i^2 = 2564.625, \quad \sum x_i y_i = 6295.$$

Zbadać, czy istnieje zależność między cechami. Wyznaczyć liniową funkcję zależności przeciętnej efektywności nawożenia łąk azotem od poziomu nawożenia.

**8.5.** Poniższe dane dotyczą ciężaru owoców ( $Y$ ) pewnej rośliny oraz ilości zastosowanego pewnego preparatu ( $X$ ) na dziesięciu poletkach:

$$\bar{x} = 3.9, \quad \bar{y} = 3.6, \quad \sum x_i^2 = 152.82, \quad \sum y_i^2 = 131.22, \quad \sum x_i y_i = 141.21.$$

Zbadać, czy istnieje zależność między cechami. Wyznaczyć liniową funkcję zależności przeciętnej ciężaru owoców od ilości zastosowanego preparatu.

**8.6.** Zbadać, czy istnieje korelacja między wielkością produkcji  $X$  pewnego artykułu (w mln metrów), a zużyciem  $Y$  pary technologicznej (w tys. ton). Dane pochodzą z dziesięciu wylosowanych zakładów.

$x_i$	1.5	3.0	2.0	3.5	1.5	4.5	2.5	4.0	4.5	3.0
$y_i$	4.5	7.0	7.5	6.5	6.5	7.5	5.5	4.5	5.5	5.0

Jeżeli zależność istnieje, to opisać ją za pomocą liniowej funkcji regresji.

**8.7.** W pewnym gospodarstwie wiejskim badano w ciągu dziesięciu kolejnych lat przeciętne dzienne spożycie ziemniaków w  $kg$  ( $X$ ) i wielkość spożycia artykułów zbożowych w  $kg$  ( $Y$ ) przypadającą na jednego członka rodziny. Z badać, czy istnieje zależność między cechami  $X$  oraz  $Y$ . Jeżeli zależność istnieje, to opisać ją za pomocą liniowej funkcji regresji.

$x_i$	0.70	0.60	0.80	0.85	0.55	0.65	0.90	1.00	0.75	0.50
$y_i$	0.50	0.70	0.50	0.40	0.75	0.60	0.30	0.20	0.55	0.70

**8.8.** Na podstawie poniższych danych zbadać, czy istnieje zależność między zawartością tłuszczu ( $X$ ) i białka ( $Y$ ) w mleku krów.

$$\sum_{i=1}^{10} x_i = 38.6, \quad \sum_{i=1}^{10} x_i^2 = 150.16, \quad \sum_{i=1}^{10} y_i = 36.2, \quad \sum_{i=1}^{10} y_i^2 = 131.74, \quad \sum_{i=1}^{10} x_i y_i = 140.34$$

Jeżeli tak, to opisać ilościowo tę zależność.

**8.9.** Na podstawie poniższych danych zbadać, czy istnieje zależność między grubością włókna lnu ( $X$ ) i grubością łądygi ( $Y$ ).

$$\sum_{i=1}^8 x_i = 83.5, \quad \sum_{i=1}^8 x_i^2 = 1269.25, \quad \sum_{i=1}^8 y_i = 8.74, \quad \sum_{i=1}^8 y_i^2 = 18.5454, \quad \sum_{i=1}^8 x_i y_i = 145.780$$

Jeżeli tak, to opisać ilościowo tę zależność.

**8.10.** Spośród studentów pewnego wydziału wylosowano niezależnie dziesięciu studentów IV roku i otrzymano dla nich następujące średnie oceny uzyskane na I roku oraz na IV roku. Z badać, czy istnieje zależność między wynikami studiów na I i na IV roku. Jeżeli taka zależność istnieje, to opisać ją za pomocą liniowej funkcji regresji.

I rok	3.5	4.0	3.8	4.6	3.9	3.0	3.5	3.9	4.5	4.1
IV rok	4.2	3.9	3.8	4.5	4.2	3.4	3.8	3.9	4.6	4.0