



№ 24-25 Дәріс

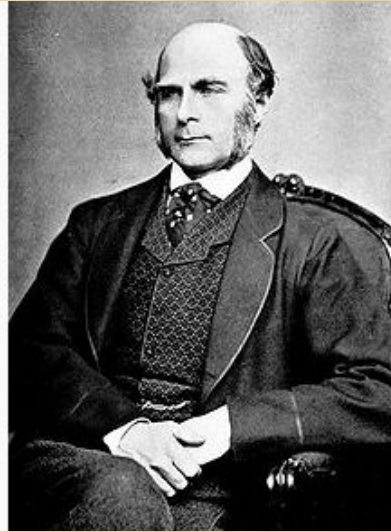
Тақырыбы:
Регрессиялық талдау

Дәріс жоспары:

1. Регрессиялық талдау: негізгі түсініктер.
2. Регрессия түрлері.
3. Сызықты регрессия параметрлерін ең кіші квадраттар әдісі бойынша бағалау.
4. Регрессия коэффициентінің маңыздылығы туралы болжамды тексеру.
5. Регрессия теңдеуінің маңыздылығы туралы болжамды тексеру.
6. Детерминация коэффициенті.
7. Жұптасқан сызықты регрессия теңдеуін тұрғызудың мысалдары және оның талдауы.

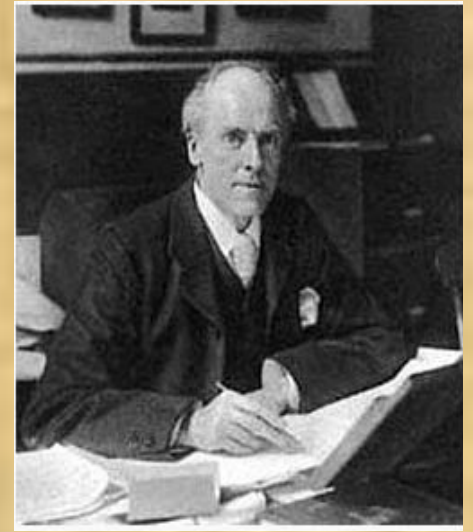
Регрессиялық талдау

Регрессиялық талдау - бір немесе бірнеше белгілердің (факторлық белгілердің) және салдардың (нәтижелі белгілердің) арасындағы байланысты өлшеуге мүмкіндік беретін статистикалық өңдеу әдісі.



Ф. Гальтон
(1822—1911)

«Регрессия» терминін алғаш рет биометрияның негізін салушы Ф. Гальтон енгізген, оның ойын ізбасары К. Пирсон дамытқан.



К. Пирсон
(1857—1936)

Регрессиялық талдау: негізгі түсініктер

Белгі – бұл зерттелетін құбылыстың немесе үдерістің негізгі ерекшелігі.

Нәтижелік белгі – зерттелетін көрсеткіш.

Факторлық белгі - нәтижелік белгінің мәніне әсер ететін көрсеткіш.

Регрессиялық талдаудың мақсаты **регрессия теңдеуі** түрінде берілген орташа мәннің (y) нәтижелік белгісінің, (x_1, x_2, \dots, x_n) , факторлық белгіге функционалдық байланысын бағалау болып табылады:

$$\bar{y} = f(x_1, x_2, \dots, x_n)$$

Регрессиялық талдау: негізгі түсініктер

РЕГРЕССИЯ

ЖҰПТАСҚАН

$$y = f(x)$$



КӨПШЕ

$$\bar{y} = f(x_1, x_2, \dots, x_n)$$

Регрессиялық талдау келесі кезеңдерден тұрады:

- функция түрін анықтаудан;
- регрессия коэффициенттерін анықтаудан;
- нәтижелі белгінің теориялық мәндерін есептеуден;
- регрессия коэффициентінің статистикалық маңыздылығын тексеруден;
- регрессия теңдеуінің статистикалық маңыздылығын тексеруден.

Дәріс жоспары:

1. Регрессиялық талдау: негізгі түсініктер.
2. Регрессия түрлері.
3. Сызықты регрессия параметрлерін ең кіші квадраттар әдісі бойынша бағалау.
4. Регрессия коэффициентінің маңыздылығы туралы болжамды тексеру.
5. Регрессия теңдеуінің маңыздылығы туралы болжамды тексеру.
6. Детерминация коэффициенті.
7. Жұптасқан сызықты регрессия теңдеуін тұрғызудың мысалдары және оның талдауы.

Регрессия түрлері

РЕГРЕССИЯ

ТУРА

тәуелсіз « x »

шамасының **артуына**
немесе кемуіне сәйкес

тәуелді « y »

шамасының **арту**
немесе кему шартына
байланысты пайда
болады

КЕРІ

тәуелсіз « x »

шамасының **артуына**
немесе кемуіне сәйкес

тәуелді « y »

шамасының **кему**
немесе арту шартына
байланысты пайда
болады

Жұптасқан регрессия теңдеуілер

Байланысты сипаттау үшін келесі жұптасқан регрессия теңдеулерінің түрлерін қолданады:

- $y=a+bx$ – **СЫЗЫҚТЫҚ**;
- $y=e^{ax+b}$ – **ЭКСПОНЕНЦИАЛДЫ**;
- $y=a+b/x$ – **ГИПЕРБОЛАЛЫҚ**;
- $y=a+b_1x+b_2x^2$ – **ПАРАБОЛАЛЫҚ**;
- $y=ab^x$ – **КӨРСЕТКІШТІК ЖӘНЕ Т.Б.**

мұндағы a, b_1, b_2 - теңдеудің коэффициенттері (параметрлері); y – нәтижелі белгі; x – факторлық белгі.

Дәріс жоспары:

1. Регрессиялық талдау: негізгі түсініктер.
2. Регрессия түрлері.
3. **Сызықты регрессия параметрлерін ең кіші квадраттар әдісі бойынша бағалау.**
4. Регрессия коэффициентінің маңыздылығы туралы болжамды тексеру.
5. Регрессия теңдеуінің маңыздылығы туралы болжамды тексеру.
6. Детерминация коэффициенті.
7. Жұптасқан сызықты регрессия теңдеуін тұрғызудың мысалдары және оның талдауы.

Сызықты регрессия параметрлерін ең кіші квадраттар әдісі бойынша бағалау

Регрессия теңдеуін құру, оның коэффициенттерін (параметрлерін) бағалауға алып келеді, ол үшін **ең кіші квадраттар әдісі** қолданылады.

Ең кіші квадраттар әдісі: $\sum (y - y_x)^2 \rightarrow \min$ болғанда параметрлерді бағалауға мүмкіндік береді.

Ең кіші квадраттар әдісі бойынша сызықты регрессия $y = a + bx$ теңдеуінің параметрлерін анықтау формуласы:

$$a = \bar{y} - b\bar{x}, \quad b = \frac{\overline{yx} - \bar{y} \cdot \bar{x}}{\overline{x^2} - \bar{x}^2}.$$

Дәріс жоспары:

1. Регрессиялық талдау: негізгі түсініктер.
2. Регрессия түрлері.
3. Сызықты регрессия параметрлерін ең кіші квадраттар әдісі бойынша бағалау.
4. Регрессия коэффициентінің маңыздылығы туралы болжамды тексеру.
5. Регрессия теңдеуінің маңыздылығы туралы болжамды тексеру.
6. Детерминация коэффициенті.
7. Жұптасқан сызықты регрессия теңдеуін тұрғызудың мысалдары және оның талдауы.

Регрессия коэффициентінің маңыздылығын тексерудің сызбасы

1) $H_0: a=0, b=0.$

$H_1: a \neq 0, b \neq 0.$

2) $p=0,05$ – маңыздылық деңгейі.

3) $t_{b \text{ есен}} = \frac{b}{m_b}, t_{a \text{ есен}} = \frac{a}{m_a}$, мұндағы m_b, m_a - кездейсоқ қателіктер:

$$m_b = \sqrt{\frac{\sum (y - y_x)^2}{n-2} \cdot \frac{1}{\sum (x - \bar{x})^2}}, \quad m_a = \sqrt{\frac{\sum (y - y_x)^2}{n-2} \cdot \frac{\sum x^2}{n \sum (x - \bar{x})^2}}.$$

4) $t_{\text{кесте}}(p; f)$, мұндағы $f=n-k-1$, n - бақылау саны, k - айнымалылары « x » теңдеудегі параметрлер саны.

5) Егер $t_{\text{есен}} > t_{\text{кесте}}$ болса, онда H_0 қабылданбайды, яғни коэффициент маңызды.

Егер $t_{\text{есен}} < t_{\text{кесте}}$ болса, онда H_0 қабылданады, яғни коэффициент маңызды емес.

Дәріс жоспары:

1. Регрессиялық талдау: негізгі түсініктер.
2. Регрессия түрлері.
3. Сызықты регрессия параметрлерін ең кіші квадраттар әдісі бойынша бағалау.
4. Регрессия коэффициентінің маңыздылығы туралы болжамды тексеру.
5. Регрессия теңдеуінің маңыздылығы туралы болжамды тексеру.
6. Детерминация коэффициенті.
7. Жұптасқан сызықты регрессия теңдеуін тұрғызудың мысалдары және оның талдауы.

Регрессия теңдеуінің маңыздылығын тексерудің сызбасы

1) H_0 : регрессия теңдеуі маңызды емес.

H_1 : регрессия теңдеуі маңызды.

2) $p=0,05$ – маңыздылық деңгейі.

$$3) F_{есеп} = \frac{\frac{\sum (y_x - \bar{y})^2}{k}}{\frac{\sum (y - y_x)^2}{n - k - 1}} = (n - 2) \frac{r_{xy}^2}{1 - r_{xy}^2},$$

мұндағы n - бақылау саны; k - айнымалылары « x » теңдеудегі параметрлер саны; y - нәтижелі белгінің нақты мәні; y_x - нәтижелі белгінің теориялық мәні; r_{xy} - жұпталған корреляция коэффициенті.

$$4) F_{кесте}(p; f_1; f_2),$$

мұндағы $f_1=k$, $f_2=n-k-1$ - еркіндік дәрежелерінің саны

5) Егер $F_{есеп} > F_{кесте}$, онда регрессия теңдеуі дұрыс таңдалған.

Егер $F_{есеп} < F_{кесте}$, онда регрессия теңдеуі дұрыс таңдалмаған.

Дәріс жоспары:

1. Регрессиялық талдау: негізгі түсініктер.
2. Регрессия түрлері.
3. Сызықты регрессия параметрлерін ең кіші квадраттар әдісі бойынша бағалау.
4. Регрессия коэффициентінің маңыздылығы туралы болжамды тексеру.
5. Регрессия теңдеуінің маңыздылығы туралы болжамды тексеру.
6. **Детерминация коэффициенті.**
7. Жұптасқан сызықты регрессия теңдеуін тұрғызудың мысалдары және оның талдауы.

Детерминация коэффициенті

Регрессиялық талдаудың сапалық өлшемінің негізгі көрсеткіші **детерминация коэффициенті (R^2)** болып табылады

Детерминация коэффициенті «у» айнымалының қандай бөлігі талдауда ескерілгендігін және талдауға енгізілген фактордың туғызатын әсерін көрсетеді.

Детерминация коэффициенті (R^2) $[0,1]$ аралығында мәндерді қабылдайды. Егер $R^2 \geq 0,8$ болса, регрессия теңдеуі сапалы болып табылады.

Детерминация коэффициенті корреляция коэффициентінің квадратына тең, яғни:

$$R^2 = r_{xy}^2.$$

Дәріс жоспары:

1. Регрессиялық талдау: негізгі түсініктер.
2. Регрессия түрлері.
3. Сызықты регрессия параметрлерін ең кіші квадраттар әдісі бойынша бағалау.
4. Регрессия коэффициентінің маңыздылығы туралы болжамды тексеру.
5. Регрессия теңдеуінің маңыздылығы туралы болжамды тексеру.
6. Детерминация коэффициенті.
7. Жұптасқан сызықты регрессия теңдеуін тұрғызудың мысалдары және оның талдауы.

Жұптасқан сызықты регрессия теңдеуін тұрғызудың мысалдары және оның талдауы

1 мысал. Төмендегі берілгендер бойынша регрессия теңдеуін тұрғызу және талдау:

1000 тұрғынға келетін тұрмаумен ауыру, x	352	228	340	300	196	258	237
1000 тұрғынға келетін пневманиямен ауыру, y	64	60	52	48	46	41	32

Шешуі.

1) Корреляция коэффициентін есептеу: $r_{xy} = 0,47$

Белгілер арасындағы байланыс түзу және қалыпты.

Жұптасқан сызықты регрессия теңдеуін тұрғызудың мысалдары және оның талдауы

1 мысал.

Шешуі (жалғасы).

2) 2.1) Есептелген кесте құру.

№	x	y	xy	x^2
1	352	64	22528	123904
2	228	60	13680	51984
3	340	52	17680	115600
4	300	48	14400	90000
5	196	46	9016	38416
6	258	41	10578	66564
7	237	32	7584	56169
Қосындысы	1911	343	95466	542637
Орташа	273	49	13638	77519,6

Жұптасқан сызықты регрессия теңдеуін тұрғызудың мысалдары және оның талдауы

1 мысал.

Шешуі (жалғасы).

2.2) Регрессия коэффициенттерін есептеу:

$$b = \frac{\overline{yx} - \bar{y} \cdot \bar{x}}{\overline{x^2} - \bar{x}^2} = \frac{13638 - 49 \cdot 273}{77519,6 - 273^2} = 0,087,$$

$$a = \bar{y} - b\bar{x} = 49 - 0,087 \cdot 273 = 25,17.$$

Жұптасқан сызықты регрессия теңдеуі: $y_x = 25,17 + 0,087x$.

Жұптасқан сызықты регрессия теңдеуін тұрғызудың мысалдары және оның талдауы

1 мысал.

Шешуі (жалғасы).

3) « x » нақты мәндерді регрессия теңдеуіне қою арқылы « y_x » теориялық мәндерін табу.

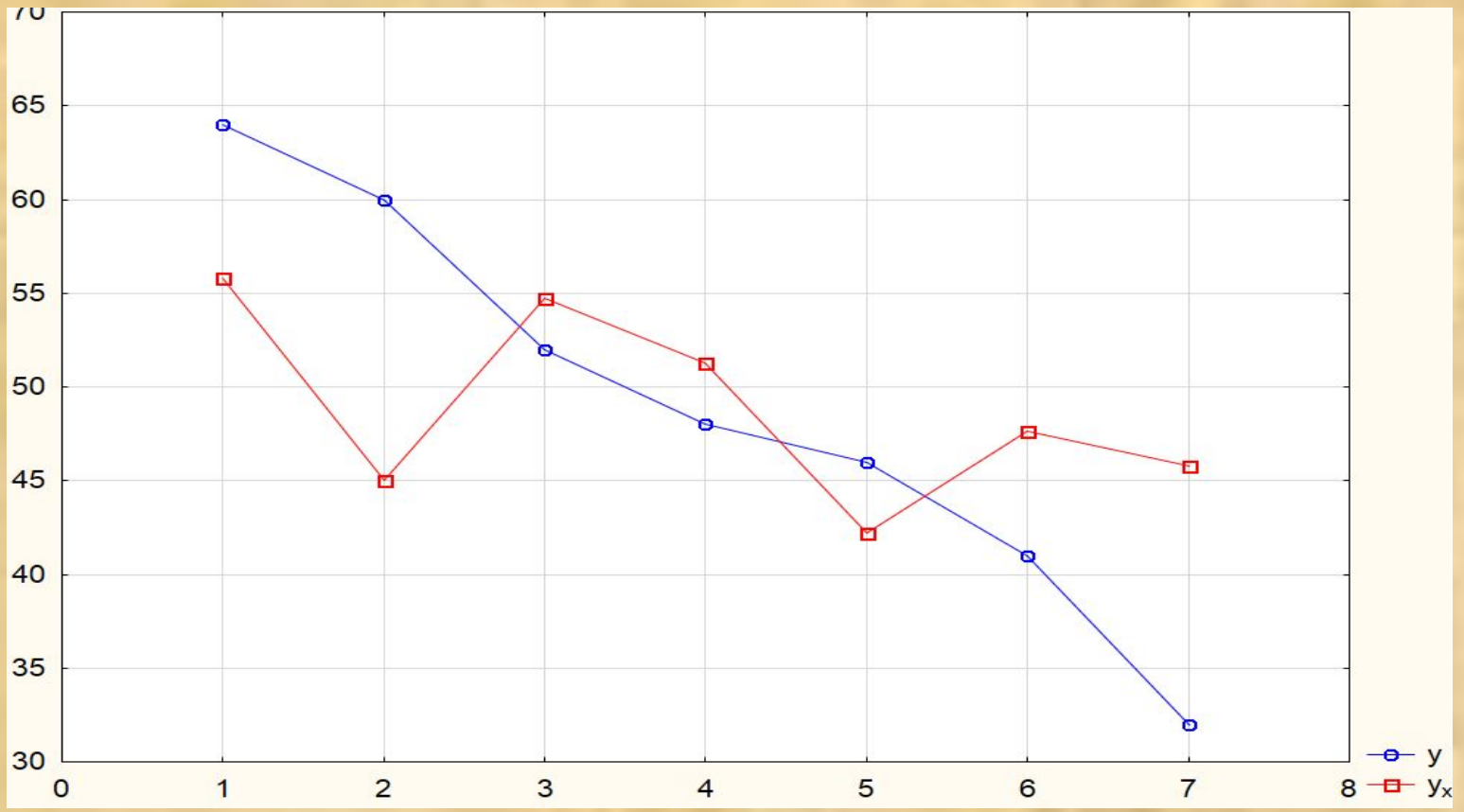
№	x	y	xy	x^2	y_x
1	352	64	22528	123904	55,89
2	228	60	13680	51984	45,07
3	340	52	17680	115600	54,85
4	300	48	14400	90000	51,36
5	196	46	9016	38416	42,28
6	258	41	10578	66564	47,69
7	237	32	7584	56169	45,86
Қосындысы	1911	343	95466	542637	343
Орташа	273	49	13638	77519,6	49

Жұптасқан сызықты регрессия теңдеуін тұрғызудың мысалдары және оның талдауы

1 мысал.

Шешуі (жалғасы).

4) Нәтижелі белгінің нақты « y » және теориялық « y_x » мәндері бойынша сызба тұрғызу .



Жұптасқан сызықты регрессия теңдеуін тұрғызудың мысалдары және оның талдауы

1 мысал.

Шешуі (жалғасы).

5) Регрессия коэффициенттерінің статистикалық маңыздылығын тексеру.

5.1) Есептелген кесте құру.

№	x	y	xy	x^2	y_x	$(x - \bar{x})^2$	$(y - y_x)^2$	
1	352	64	22528	123904	55,89	47,54	65,70	
2	228	60	13680	51984	45,07	15,42	222,83	
3	340	52	17680	115600	54,85	34,19	8,11	
4	300	48	14400	90000	51,36	5,55	11,27	
5	196	46	9016	38416	42,28	45,16	13,84	
6	258	41	10578	66564	47,69	1,71	44,77	
7	237	32	7584	56169	45,86	9,87	192,05	
Қосын- дысы	1911	343	95466	542637	343	159,45	558,55	
Орташа	273	49	13638	77519,6	49	22,78	79,79	

Жұптасқан сызықты регрессия теңдеуін тұрғызудың мысалдары және оның талдауы

1 мысал.

Шешуі (жалғасы).

5.2) Кездейсоқ қателіктерді есептеу:

$$m_b = \sqrt{\frac{\sum (y - y_x)^2}{n - 2}} = \sqrt{\frac{558,55}{7 - 2}} \approx 0,073,$$

$$m_a = \sqrt{\frac{\sum (y - y_x)^2}{n - 2} \cdot \frac{\sum x^2}{n \sum (x - \bar{x})^2}} = \sqrt{\frac{79,8}{7 - 2} \cdot \frac{542637}{7 \cdot 2990,6}} \approx 20,34.$$

Жұптасқан сызықты регрессия теңдеуін тұрғызудың мысалдары және оның талдауы

1 мысал.

Шешуі (жалғасы).

$$5.3) \quad t_{b \text{ есен}} = \frac{b}{m_b} = \frac{0,087}{0,073} \approx 1,19,$$

$$t_{a \text{ есен}} = \frac{a}{m_a} = \frac{25,17}{20,34} \approx 1,24.$$

$$5.4) \quad t_{\text{кесте}}(0,05; 5) = 2,57.$$

$$5.5) \quad \begin{aligned} t_{b \text{ есен}} &< t_{\text{кесте}}, \text{ яғни «}b\text{» коэффициенті – маңызды емес,} \\ t_{a \text{ есен}} &< t_{\text{кесте}}, \text{ яғни «}a\text{» коэффициенті – маңызды емес.} \end{aligned}$$

Жұптасқан сызықты регрессия теңдеуін тұрғызудың мысалдары және оның талдауы

1 мысал.

Шешуі (жалғасы).

б) Регрессия теңдеуінің статистикалық маңыздылығын тексеру:

$$6.1) F_{есеп} = \frac{\sum (y_x - \bar{y})^2}{\sum (y - y_x)^2} = \frac{159,45}{558,55} = 1,43.$$

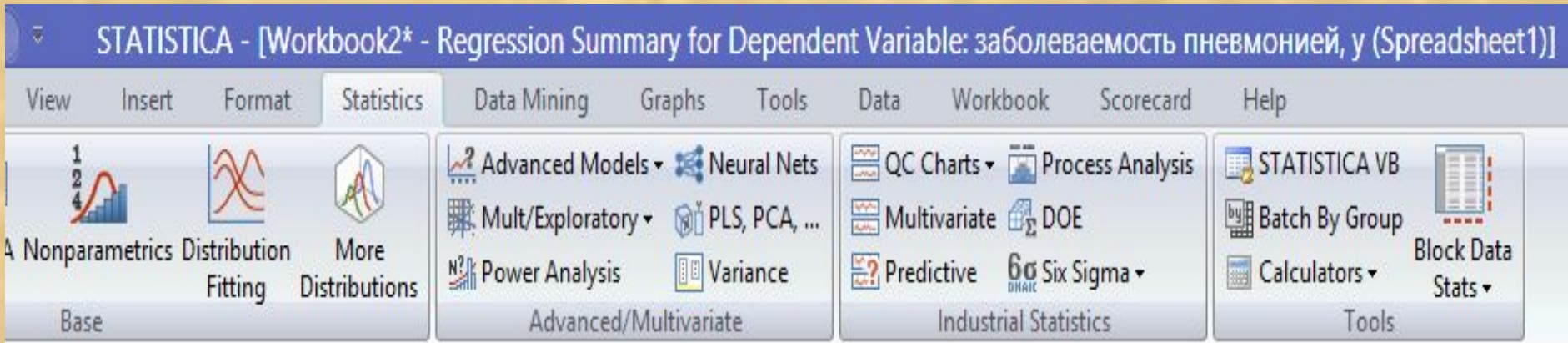
$$6.2) F_{кесте}(p; k; n-k-1) = (0,05; 1; 5) = 6,61.$$

6.3) $F_{есеп} < F_{кесте}$, онда регрессия теңдеуі дұрыс таңдап алынбаған. Бұл нәтижені орташа тәуелділік ($r_{xy} = 0,47$) және бақылау санының аз болуымен түсіндіруге болады..

Жұптасқан сызықты регрессия теңдеуін тұрғызудың мысалдары және оның талдауы

1 мысал.

Шешуі (жалғасы).



Regression Summary for Dependent Variable: заболеваемость пневмонией, y						
R= ,47124875 R ² = ,22207539 Adjusted R ² = ,06649047						
F(1,5)=1,4274 p<,28577 Std.Error of estimate: 10,569						
	коэффициент корреляции	стандарт. ошибка коэффициента корреляции	коэффициенты регрессии	стандарт. ошибки коэффициентов регрессии	расчетные значения t-критерия Стьюдента	p-value
N=7						
коэффициент а			25,17412	20,33881	1,237738	0,270763
коэффициент b	0,471249	0,394443	0,08727	0,07305	1,194721	0,285772

Бақылау сұрақтары:

1. Регрессиялық талдау қандай кезеңдерден тұрады?
2. Регрессияның қандай түрлерін білесіңдер?
3. Жұптасқан сызықтық регрессия теңдеуінің коэффициенттері қалай анықталады?
4. Регрессия коэффициентінің маңыздылығы қалай тексеріледі?
5. Регрессия теңдеуінің маңыздылығы қалай тексеріледі?

