Сызықтық емес (бейсызықтық) жұп регрессия

Біз осы уақытқа дейін тек сызықтық жұп регрессия қарастырған болатынбыз. Оның теңдеуінің түрін еске сала кетейік:

$$y = a + bx$$

Егер тәуелсіз фактор мен қорытынды фактордың арасындағы байланыс сызықтық болмайтын болса, онда олардың регрессиялық моделі сызықтық емес (бейсызықтық) функция арқылы өрнектеледі. Бұндай регрессияны *<u>сызықтық емес</u>*, кейде бейсызықтық регрессия деп атаймыз.

Қисық сызықты регрессияны екі түрге бөліп қарастыруға болады:

1. тәуелсіз факторлар бойынша бейсызықтық, бірақ параметрлері сызықтық түрде теңдеуге кіреді. Мысалы: параболалар, гиперболалар, логарифмдік функциялар;

$$y = a + bx + cx^2; \qquad y = a + \frac{b}{x}; \qquad y = a + \ln x.$$

2.тек қана тәуелсіз факторлар ғана емес параметрлері бойынша да сызықтық емес регрессиялар. Мысалы: дәрежелік, көрсеткіштік, т.с.с. функциялар. $y = a + b^x; \qquad y = a + x^b.$

Іс жүзінде жиі кездесетін сызықтық емес жұп регрессияның бірі көпмүшелік болған кез:

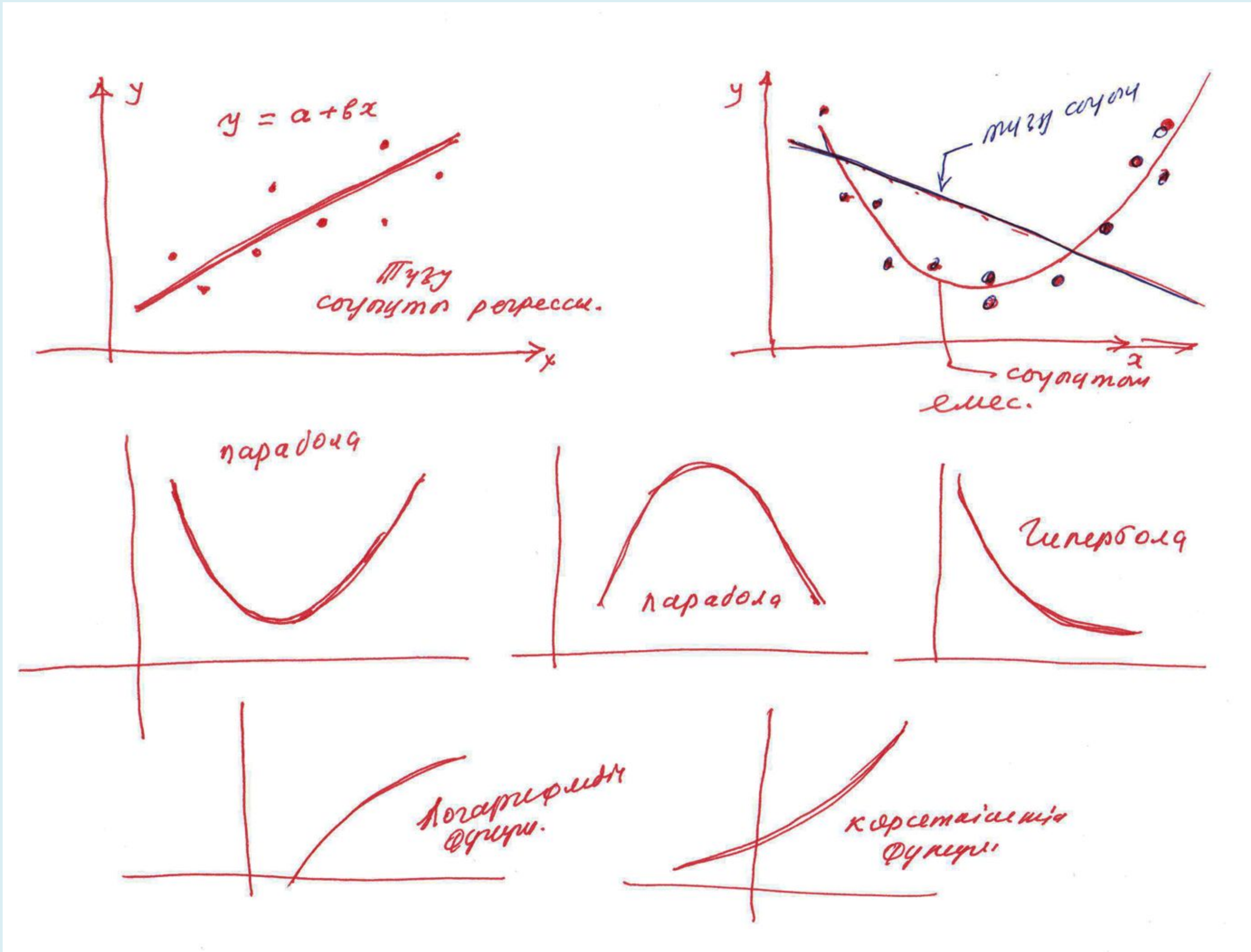
$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + ... + a_k x^k$$

мұнда a_0, a_1,… $a_к$ белгісіз коэффициенттер, олар төмендегі теңдеулер жүйесінін шешімі болады:

$$\begin{cases} a_0 n + a_1 \sum x_i + a_2 \sum x_i^2 + ... + a_k \sum x_i^k = \sum y_i; \\ a_0 \sum x_i + a_1 \sum x_i^2 + a_2 \sum x_i^3 + ... + a_k \sum x_i^{k+1} = \sum x_i y_i; \\ .... \\ a_0 \sum x_i^k + a_1 \sum x_i^{k+1} + a_2 \sum x_i^{k+2} + ... + a_k \sum x_i^{2k} = \sum x_i^k y_i. \end{cases}$$

Бұл теңдеулер жүйесін Крамер, немесе Гаусс әдістерінің бірімен шешіп белгісіз коэффициенттерді a_0, a_1,… $a_к$ табуға болады.

Егер регрессия теңдеуі көпмүшелік болмаса басқаша әдіспен жұмыс істеуге тура келеді. Көп жағдайда сызықтық емес регрессия теңдеулерінің параметрлерін сызықтық регрессия теңдеуінің параметрлерін тапқан әдіске келтіру арқылы шығарады.


y
y = a + bx
x
y
x
парабола
парабола
Гипербола

Әңгімеміз түсінікті болуы үшін сызықтық емес регрессия теңдеуін $y=1/(a_0+a_1x)$ түрінде іздестіргенде қолданатын әдісті толығырақ көрсетелік:

Бізге тәжірибеден алынған мынандай таблица беріледі (сол жақтағы таблица). Енді осы мәліметтерді қолданып келтірілген сызықтық емес регрессия теңдеуінің параметрлерін іздестірелік.

x	y
x_1	y_1
x_2	y_2
…	…
x_n	y_n

Ол үшін мынандай түрлендіру жасаймыз: $\tilde{y}=1/y$. Сонда теңдеу мына түрге келеді:

$$\frac{1}{\tilde{y}} = \frac{1}{a_0 + a_1 x} \Rightarrow \tilde{y} = a_0 + a_1 x$$

x	y	$\tilde{y}=1/y$
x_1	y_1	$\tilde{y}_1$
x_2	y_2	$\tilde{y}_1$
…	…	…
x_n	y_n	$\tilde{y}_n$

Ал соңғы теңдеу кәдімгі өзімізге белгілі сызықтық регрессия теңдеуі. Түрлендіру формуласының көмегі арқылы жаңа таблица (оң жақтағы таблица) құрамыз да, бұрыннан белгілі әдіспен a_0, a_1 параметрлерін табамыз. Бұл параметрлер табылғасын сызықтық емес регрессия теңдеуін жаза аламыз.

Осы тәрізді әдісті қолданып көп жағдайда сызықтық емес регрессия теңдеуінің параметрлерін табуға болады.

Төменде сызықтық емес регрессия теңдеуінің түрлері, қолданылатын түрлендірулер, түрлендіру қолданғаннан кейін шыққан сызықтық регрессия теңдеуінің түрлері келтірілген.

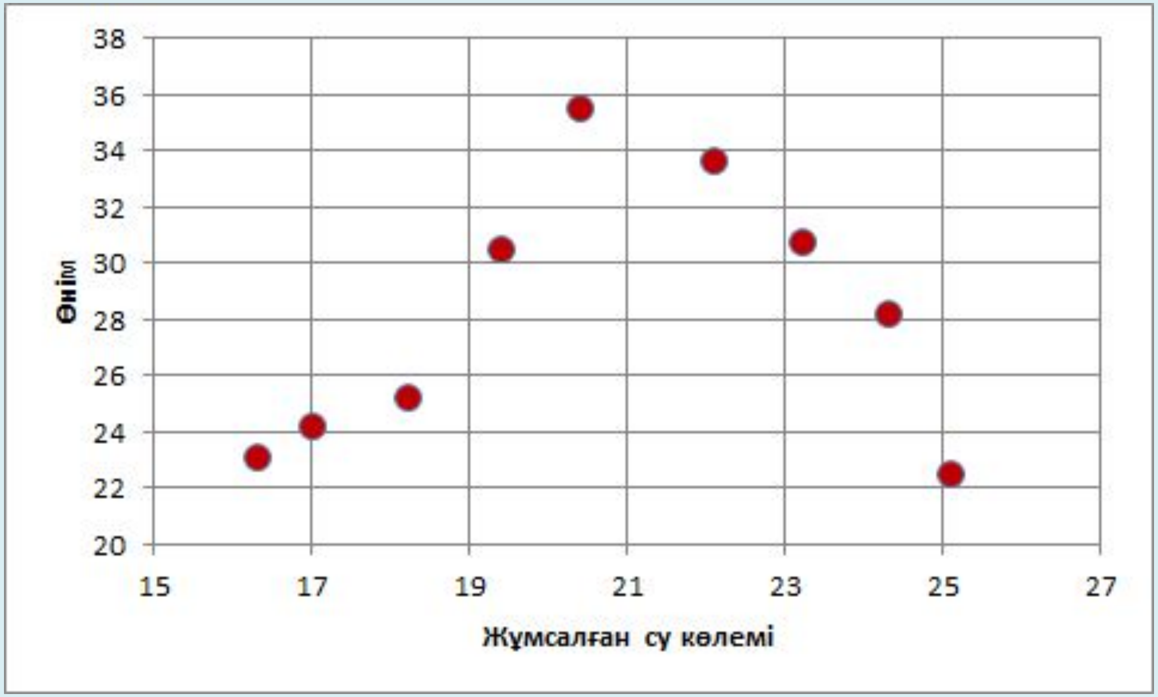
Қисық сызықты регрессия түрі	Қолданылатын түрлендірулер	Сызықтық түрге келтірілген түрі
$y = \frac{1}{a_0 + a_1 x}$	$\tilde{y} = \frac{1}{y}$	$\tilde{y} = a_0 + a_1 x$
$y = a_0 + \frac{a_1}{x}$	$\tilde{x} = \frac{1}{x}$	$y = a_0 + a_1 \tilde{x}$
$y = \frac{x}{a_0 + a_1 x}$	$\tilde{y} = \frac{x}{y}$	$\tilde{y} = a_0 + a_1 x$
$y = \frac{1}{a_0 + a_1 e^{-x}}$	$\tilde{x} = e^{-x}$, $\tilde{y} = \frac{1}{y}$	$\tilde{y} = a_0 + a_1 \tilde{x}$
$y = a_0 e^{a_1 x}$	$\tilde{y} = \ln x$	$\tilde{y} = \ln a_0 + a_1 \tilde{x}$
$y = \sqrt{a_0 + a_1 x}$	$\tilde{y} = y^2$	$\tilde{y} = a_0 + a_1 x$
$y = e^{a_0 + a_1 e^x}$	$\tilde{x} = e^x$, $\tilde{y} = \ln y$	$\tilde{y} = a_0 + a_1 \tilde{x}$
$y = a_0 + a_1 \log x$	$\tilde{x} = \log x$	$y = a_0 + a_1 \tilde{x}$
$y = \frac{a_0 x}{a_1 + x}$	$\tilde{x} = \frac{1}{x}$, $\tilde{y} = \frac{1}{y}$	$\tilde{y} = \frac{1}{a_0} + \frac{a_1}{a_0} \tilde{x}$

Параболалық регрессияға мысал

Капуста өсіру кезінде суаруға жұмсалған су көлемі мен капуста өнімділігінің арасындаы байланысты табу керек болсын. Байқаудан мынадай мәліметтер алынған.

	18,2	16,3	17,0	19,4	20,4	22,1	23,2	24,3	25,1
	25,3	23,1	24,2	30,5	35,6	33,7	30,8	28,2	22,5

Регрессия теңдеуін табыңыз



$$\begin{cases} 9\cdot a_0 + 186\cdot a_1 + 3925{,}6\cdot a_2 = 253{,}9; \\ 186\cdot a_0 + 3925{,}6\cdot a_1 + 84506{,}55\cdot a_2 = 5275{,}67; \\ 3925{,}6\cdot a_0 + 84506{,}55\cdot a_1 + 1852505{,}77\cdot a_2 = 111670{,}14. \end{cases}$$

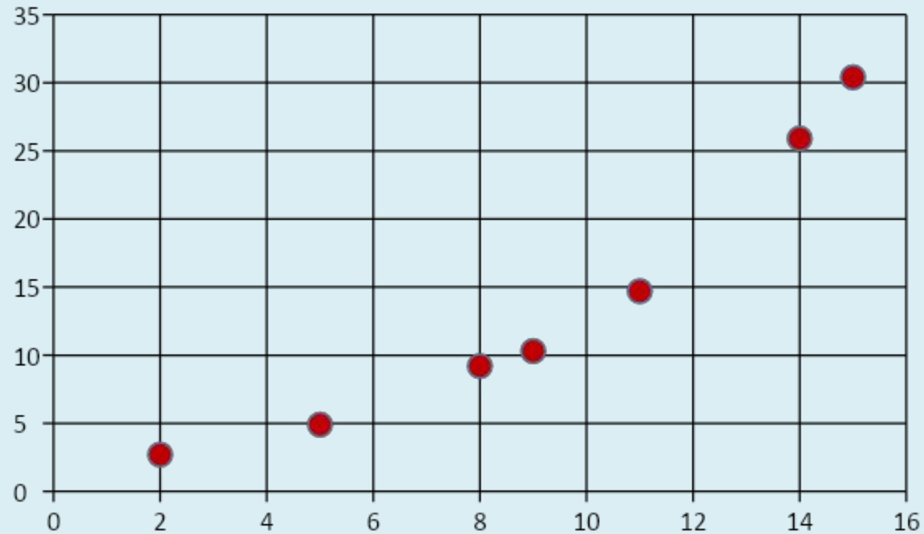
$$y = -212{,}96 + 23{,}45x - 0{,}56x^2.$$

Енді *$y=ab^x$* түріндегі регрессия теңдеуінің параметрлерін табудыңтмысалын қарастыралық. Тәжірибеден алынған төмендегі мәліметтерді қарастыралық:

	9	*8*	*11*	*2*	*5*	*14*	*15*
	10,3	*9,2*	*14,7*	*2,7*	*4,9*	*25,9*	*30,4*

Бізге регрессия теңдеуін табу керек болсын дейік.

Әуелі регрессия теңдеуінің түрін анықтау керек. Ол үшін корреляция өрісін саламыз. Коррреляция өрісіндегі нүктелердің орналасы ретіне қарап біз регрессия теңдеуінің түрі *$y=ab^x$* болады деп болжауымызға болдыю



y=ab^x **теңдеуінің екі жағын логарифмдесек *lny=lna+(lnb)x* өрнегін аламыз. Енді мынандай белгілеулер енгізелік:**

$$\tilde{y} = \ln y; \qquad \tilde{a} = \ln a; \qquad \tilde{b} = \ln b.$$

Сонда іздеп отырған теңдеуіміз мына түрге келеді:

$$\tilde{y} = \tilde{a} + \tilde{b}x$$

Ал бұл бізге бұрыннан таныс сызықтық регрессия. Біз жаңа таблица құрамыз:

x	9.0	8.0	11.0	2.0	5.0	14.0	15.0
y	10.3	9.2	14.7	2.7	4.9	25.9	30.4
ỹ=lny	2.3	2.2	2.7	1.0	1.6	3.3	3.4

Бұрыннан белгілі әдіспен сызықтық регрессия теңдеуінің параметрлерін табамыз:

$$\tilde{a} = 0{,}6615; \qquad \tilde{b} = 0{,}1853 \qquad \Rightarrow \qquad a = 1{,}9377; \qquad b = 1{,}2036.$$

$$y = (1{,}9377) \cdot (1{,}2036)^x.$$