

1-LABORATORIYA ISHI

Mavzu: Eng kichik kvadratlar usuli bo'yicha bitta bog'liq bo'lmagan o'zgaruvchili regressiya. Regressiya tenglamasini interpretatsiyasi. Baholash sifati. Determinatsiya koeffitsienti.

Kerakli texnik vositalar:

Pentium-4 shaxsiy kompyuteri.

Kerakli dasturiy vositalar:

Microsoft EXCEL dasturi.

Ishning maqsadi: Microsoft EXCEL dasturida eng kichik kvadratlar usuli bo'yicha bir omilli regressiya tenglamasini hosil qilish. Regressiya tenglamasini interpretatsiyasi. Baholash sifati. Determinatsiya koeffitsientlarini aniqlash.

1-masala. Quyidagi berilgan dastlabki ma'lumotlar asosida avvalo x_1 , y keyin x_2 va y o'zgaruvchilar orasidagi regressiya koeffitsientini aniqlang, qaysi omil y ga ko'proq ta'sir ko'rsatishini aniqlang va ular haqqoniyligini o'rnatib:

jadval

Dalalar №	Kuzgi bug'doy hosildorligi, s./ga, y	Yer sifatining bali, X_1	Suv sarfi m^3/ga X_2
1	18,1+k	25	7,0
2	21,1+k	27	7,0
3	22,9+k	26	6,4
4	18,9+k	28	7,8
5	18,6+k	30	7,0
6	30,5+k	35	7,5
7	23,4+k	40	4,8
8	27,6+k	45	2,6
9	20,9+k	50	5,2
10	18,2+k	52	4,4
11	18,9+k	54	9,1

12	25,9+k	55	9,0
13	18,5+k	60	9,3
14	24,0+k	65	8,2
15	17,4+k	70	9,1
16	23,9+k	75	3,2
17	23,8+k	80	6,4
18	20,4+k	85	2,5
19	29,2+k	90	3,1
20	23,5+k	95	7,5

Bu yerda k-talabani jurnali bo'yicha tartib raqami

Yuqoridagi berilgan ma'lumotlardan foydalanib quyidagilarni bajaring: a) y – kuzgi bug'doy hosildorligi va x_1 -yer sifatining bali; b) x_2 - kuzgi bug'doy hosildorligi va suv sarfi orasidagi regressiya bog'lanishi tuzilsin. O'xshatishning sifati baholansin.

a) haqiqiy model quyidagicha ifodalansin

$$y = \alpha + \beta x + \varepsilon.$$

a) tanlama kuzatishlaridan keyin (a, b) lar baholansin.

Nazariy qism

Qo'yilgan masalani yechishning usullaridan biri xatolar kvadratlarining yig'indisini minimallashtirishdan iboratdir.

$$S = e_1^2 + e_2^2 + e_3^2 + e_4^2 \rightarrow \min$$

bundan $S(a, b) = \sum_i e_i^2 \rightarrow \min$

kelib chiqadi. Bundan quyidagini yozishimiz

mumkin: $S(a, b) = \sum_i (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - a - bx_i)^2 \rightarrow \min$

Bizga oliy matematikadan ma'lumki biror bir funksiyaning ekstremal nuqtalarini topish uchun uning birinchi tartibli hosilasi nolga tenglashtiriladi:

$$\frac{\partial S}{\partial a} = 0, \begin{cases} \frac{\partial S}{\partial a} = -2 \sum_{i=2}^n (y_i - a - bx_i) = 0 \\ \frac{\partial S}{\partial b} = -2 \sum_{i=1}^n (y_i - a - bx_i)x_i = 0 \end{cases}$$

Bu sistemada qavslarni ochib, o'xshash hadlarni ixchamlashtirganda quyidagi tenglamalar sistemasi hosil qilinadi:

$$\begin{cases} \sum y_i = na + b \sum x_i \\ \sum x_i y_i = a \sum x_i + b \sum x_i^2 \end{cases}$$

Bu tenglamalar sistemasidagi $\sum y_i, \sum x_i, \sum x_i y_i, \sum x_i^2$ yig'indilarni topib,

tenglamalar sistemasini a, b noma'lumlarga nisbatan yechilganda a va b noma'lumlarni topish mumkin yoki bu noma'lumlarni quyidagi formulalar orqali ham aniqlash mumkin:

$$b = \frac{\text{cov}(x, y)}{\text{var}(x)}; \quad a = \bar{y} - b\bar{x};$$

bu yerda

$$\text{cov}(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y});$$

$$\text{var}(x) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

Uslubiy ko'rsatma

1-masala. Quyidagi ma'lumotlar asosida: a) y – oziq-ovqatga sarf xarajatlar va x – shaxsiy daromat orasidagi; b) y oziq-ovqatga sarf xarajatlar va t – vaqt orasidagi regressiya bo'g'lanishini eng kichik kvadratlar usuli asosida hosil qilinsin(sh.b.)

Yillar	1990	1991	1992	1993	1994
--------	------	------	------	------	------

x	2	6	10	14	18
y	1	2	4	11	12

O'xshatishning sifati quyidagicha baholanadi:

a) Haqiqiy model quyidagicha ifodalansin

$$y = \alpha + \beta x + \varepsilon.$$

b) Tanlama kuzatishlaridan keyin (a,b) lar baholansin.

Dastlabki va hisoblangan ma'lumotlar jadvalda keltirilgan:

<i>Yillar</i>	x	y	x^2	xy	\hat{y}	$(y - \bar{y})^2$	$(\hat{y} - \bar{y})^2$	$(y - \hat{y})^2$
1990	2	1	4	2	-0,2	25	38,44	1,44
1991	6	2	36	12	2,9	16	9,61	0,81
1992	10	11	100	40	6	4	0	4
1993	14	12	196	154	9,1	25	9,61	3,61
1994	18	30	324	216	12,2	36	38,44	0,04
Jami	50	30	660	424	30	106	96,1	9,9
O'rtacha	10	6	132	84,8	6	21,2	19,22	1,98
	\bar{x}	\bar{y}	$\overline{x^2}$	\overline{xy}	$\bar{\hat{y}}$	$Var(y)$	$Var(\hat{y})$	$Var(e)$

Natijada quyidagiga ega bo'lamiz:

$$-60=24,8,$$

$$cov(x, y) = \overline{xy} - \bar{x}\bar{y} = 84,8$$

$$var(x) = \overline{x^2} - (\bar{x})^2 = 132 - 100 = 32,$$

$$b = \frac{cov(x, y)}{var(y)} = \frac{24,8}{32} = 0,775, \quad a = \bar{y} - b\bar{x} = 6 - 0,775 * 10 = -1,75.$$

Bundan kelib chiqadi.

$$\hat{y} = -1,75 + 0,775x$$

$b=0,775$ koeffitsient shuni ko'rsatadiki, daromadni bir birlikka oshirsak oziq-ovqatga sarf xarajatlar birligi o'rtacha 0,775 sh.b. ga oshadi.

Eslatma.Excelda (a,b) ni baholashni quyidagi funksiyalar orqali aniqlash ham mumkin:

$$a = \text{ОТРЕЗОК}(\text{массиву}_1 \text{ массивх}),$$

$$b = \text{НАКЛОН}(\text{массиву}_1 \text{ массивх}).$$

shart bajariladi.

$$\text{var}(y) = \text{var}(\bar{y}) + \text{var}(e)$$

O'xshatishning sifatini determinatsiya koeffitsienti orqali baholaymiz:

$$R^2 = \frac{\text{var}(\hat{y})}{\text{var}(y)} = \frac{19,22}{21,2} = 0,907,$$

ya'ni 90,7% erksiz o'zgaruvchining variatsiyasi regressiya orqali tushuntiriladi.

2-masala. Regressiya modelida, $Y = \beta X + \varepsilon$ ozod hadsiz β uchun eng

kichik kvadratlar usuli orqali baholash β ga teng. Bu model uchun

$$b = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{\overline{xy}}{\overline{x^2}}$$

tanlama regressiya $\hat{y} = bx$ ga teng. Erksizo'zgaruvchining kuzatuv qiymatlari

hisoblangan tenglamadan bog'liq. B ni bahosini miqdorni $y_i = \hat{y}_i + e_i$

minimallashtirish orqali topamiz.

$$Q = \sum e_i^2 = \sum (y_i - bx_i)^2 = \sum y_i^2 - 2b \sum x_i y_i + b^2 \sum x_i^2.$$

ni hosil qilamiz.

$$Q'_b = -2 \sum x_i y_i + 2b \sum x_i^2 = 0$$

Bundan

$$b = \frac{\sum x_i y_i}{\sum x_i^2} \text{ hosil bo'ladi.}$$

Yuqoridagi topshiriqdagi berilgan ma'lumotlar variantlardan foydalanib, formulaning to'g'riligini ko'rsating.

$$R^2$$

Hosildorlik va yerning sifati orasidagi bog'lanish tenglamasini quyidagi chiziqli tenglama ko'rinishida qidiramiz: $Y = a + bx$

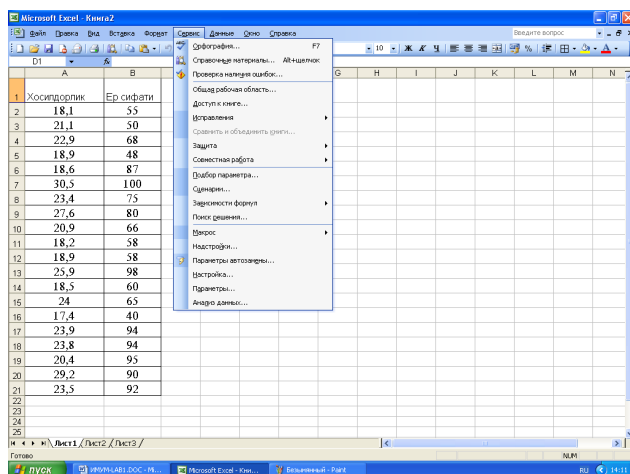
bu yerda a, b lar noma'lum parametrlar bo'lib, bu noma'lum parametrlarni eng kichik kvadratlar usulidan foydalanib baholash mumkin.

Bu noma'lumlarni topib, regressiya tenglamasini, MS EXCEL elektron jadvalida quyidagi ketma ketlikda, hosil qilish mumkin:

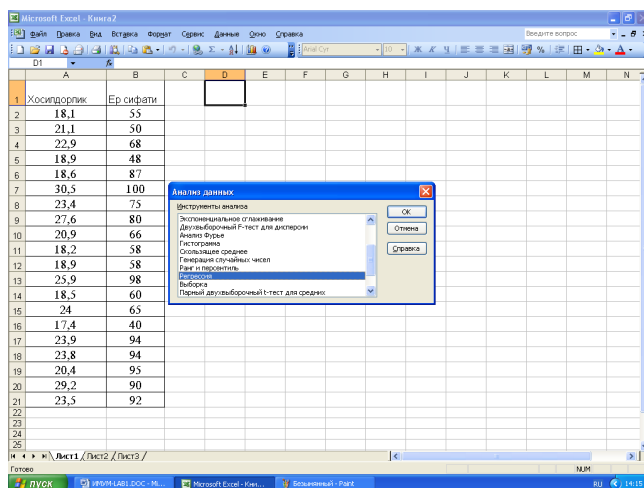
1. MS EXCEL elektron jadvaliga ma'lumotlarni kiritish kerak.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
	Хосилдорлик	Ер сифати												
2	18.1	55												
3	21.1	50												
4	22.9	68												
5	18.9	48												
6	18.6	87												
7	30.5	100												
8	23.4	75												
9	27.6	80												
10	20.9	66												
11	18.2	58												
12	18.9	58												
13	25.9	98												
14	18.5	60												
15	24	65												
16	17.4	40												
17	23.9	94												
18	23.8	94												
19	20.4	95												
20	29.2	90												
21	23.5	92												
22														
23														
24														
25														

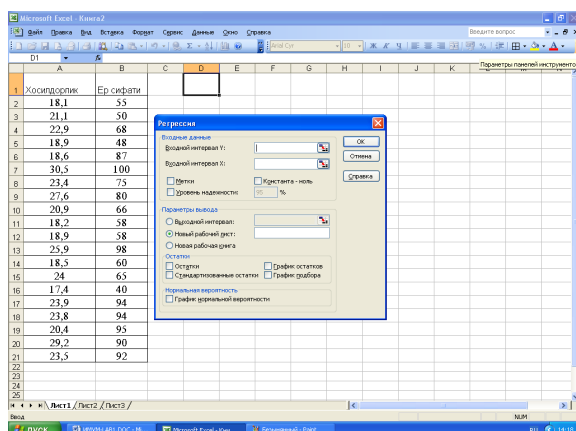
1. Kursorni D1 katagiga qo'yib, menyu qatoridan *Справка* → bo'limini tanlaganimizda quidagi muloqot oynasi chiqadi.




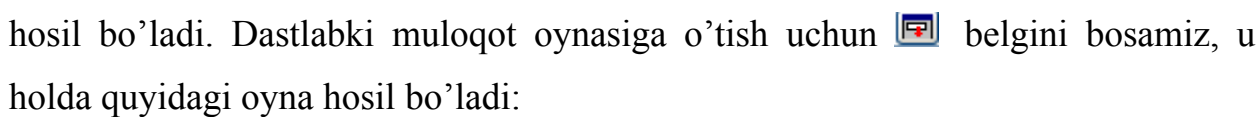
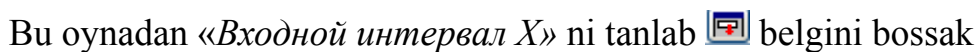
Bu yerdan «Анализ данных» bo'limini tanlaganimizda quyidagi muloqot oynasi chiqadi. Agar *Анализ данных* bo'lmasa uni «Настройка» orqali hosil qilish mumkin:



Bu muloqot oynasi ichidan *Регрессия* ni tanlab OK tugmasini bossak quyidagi muloqot oynasi hosil bo'ladi:



Muloqot oynasidagi «Входной интервал Y» to'g'risidagi  belgini bossak va «Хосилдорлик» ustunini belgilasak quyidagi hosil bo'ladi:



Наблюдения (Kuzatishlar)	20
--------------------------	----

Determinatsiya koeffitsienti

Bu qiymatlar hosildorlikning yerning sifatiga bog'liq ekanligini bildiradi.

5- jadval

	Df	Ss	Ms	F	F haqqoniyligi
Regressiya	1	134.983	134.983	16.3789	0.0007
Qoldiq	18	148.342	8.2412		
Jami	19	283.325			

Dispersiya tahlili

5-jadval regressiya, qoldiq va jamilar uchun hisoblanayotgan ko'rsatkichlarning shartli belgilaridan iborat:

- Df – erkinlik (bog'lik bo'lmagan qiymatlar darajasi soni);
- SS –chetlanishlar yig'ndisi kvadratlari soni;
- MS dispersiya, u SS/ Df nisbat orqali hisoblanadi;
- F- regressiya dispersiyasining qoldiq dispersiyaga nisbati;
- F qiymati- haqqoniylik darajasi, MS Регрессия/ MSОстаток nisbat

kabi hisoblanadi. Agar F ifoda 1 ga yaqin bo'lsa regressiya tenglamasi prognozlash uchun ahamiyatli.

6- jadval

	Y-kesishuvi	Yerning sifati, ball
Koeffitsientlar	12.0094	0.139
Standart xato	2.618	0.034
RF statistikasi	4.585	4.047
P qiymati	0.0002	0.0007
95% dan past	6.507	0.067
95% dan yuqori	17.511	0.2119

Regressiya tenglamasi parametrlari

Agar hisoblangan miqdor 1 ga yaqin bo'lsa, o'zgaruvchi koeffitsienti regressiya tenglamasida qo'llaniladi. «Hosildorlik» va «Y – kesishuvi» (regressiya tenglamasi ozod hadi) muhim hisoblanmaydi. Shuning uchun regressiya modeli tenglamasi: $Y = 12 + 0,139 * X_{ball}$ ni quyidagi ko'rinishda berish mumkin: $Y = 0,139 * X_{ball}$.