MEXANIZATSIYALASH MUHANDISLARI INSTITUTI

«Ekonometrika» fanidan

Bir o’zgaruvchili regressiya va korrelyatsiya modellari

mavzusi bo’yicha videoma’ruza

Ma’ruzachi: TIQXMMI
«Axborot texnologiyalari» kafedrasi
professori, i.f.n.
Shadmanova Gulchera

РЕЖА

1. Ekonometrik tahlilning asosiy masalalari.
2. Iqtisodiy ma'lumotlardagi chiziqli statistik bog’lanish tahlili.
3. Korrelyasiya koeffitsienti
4. Chiziqli regressiya tahlili
5. Parametrlarni baholashning eng kichik kvadratlar usuli унга доир мисол

1. Ekonometrik tahlilning asosiy masalalari.

Ekonometrik bo’lmagan modelda, X va Y o’zgaruvchilar uchun bog’lanish ko’rinishi

$$F(x,y)=0 \qquad (1)$$

bo’ladi. Ekonometrikaning asosiy masalasi (1) farazning to’g’riligini tekshirishdir. Shuning uchun ekonometrika avvalo X va Y bog’lanishini quyidagi ko’rinishda:

$$F(x,y,e)=0 \qquad (2)$$

qaraydi, bu yerda e- tasodifiy miqdor bo’lib, u ehtimollik qonuniga bo’ysinadi.

2. **Iqtisodiy ma'lumotlardagi chiziqli statistik bog’lanish tahlili.**

Iqtisodiy izlanishlardagi asosiy masalalardan biri o'zgaruvchilar orasidagi bog’lanishni tahlil qilishdir. O'zgaruvchilar orasidagi bog’lanishning eng oddiy turi chiziqli bog’lanish bo’lib, mana shu bog‘lanishning tarkibini, uning parametrlarini baholash matematik statistikaning asosiy yo'nalishlaridan biridir.

Iqtisodiy ma'lumotlardagi chiziqli statistik bog’lanish tahlili.

Ikkita X, Y o'zgaruvchilar orasidagi bog'lanish masalasini va ular orasidagi quyidagi ikkita munosabatni qaraymiz. 1. X va Y o‘zgaruvchi o'zaro chiziqli bog'lanishga egami? 2. X va Y o'zgaruvchilarning bog'lanish formulasi qanaqa? Birinchi holda X va Y ikkalasi teng xuquqli bo'lib, ular orasida erkli va erksiz o'zgaruvchisi bo'lmaydi. Bu holda korrelyatsiya tahlili o‘tkaziladi. Ikkinchi holda esa, bir o‘zgaruvchining ikkinchisiga bog‘lanishining bog‘lanish formulasi talab qilinib, bu hol uchun regressiya tahlili o'tkaziladi.

3. Korrelyasiya koeffitsienti

Chiziqli bog’lanish darajasining o‘lchami sifatida korrelyatsiya koeffitsienti ishlatiladi. X va Y o‘zgaruvchilar orasidagi korrelyatsiya koeffitsienti

$$r(x,y)=\frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2}} \quad (1)$$

ko‘rinishda ifodalanadi.

Korrelyasiya koeffitsienti

Uning qiymati –1 va +1 orasida o‘zgaradi. –1 qiymatni chiziqli manfiy bog‘lanish natijasida +1 qiymatni chiziqli musbat bog‘lanish natijasida qabul qiladi.

Korrelyatsiya koeffitsientining 0 ga yaqin qiymati o‘zgaruvchilar orasida bog‘lanish yo‘qligini bildiradi.

(1) formula bilan ishlash uchun quyidagi jadval to‘ldiriladi:

№	x	y	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x}) * (y - \bar{y}$	$(x - \bar{x})^2$	$(y - \bar{y})^2$
1	2	1	-8	-5	40	64	25
2	6	2	6	-4	-24	36	16
3	10	4	10	-2	-20	100	4
4	14	11	14	5	70	196	25
5	18	12	18	6	108	324	36
Jami	50	30			174	720	106
O‘rtacha	10	6			40	64	25

R=F7/КОРЕНЬ(G7*H7)=0,63

4. Chiziqli regressiya tahlili

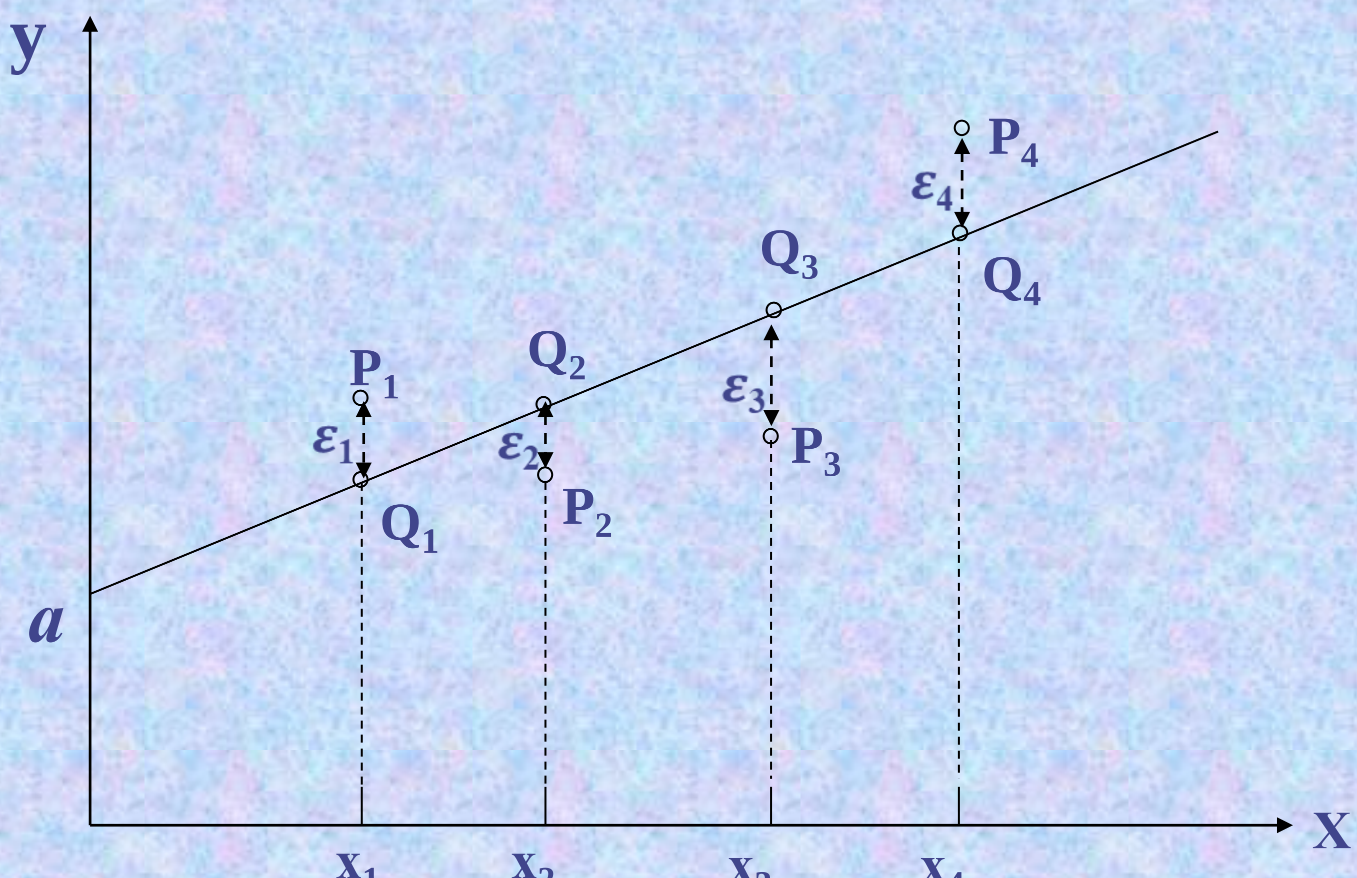
Aytaylik ikki o’zgaruvchi orasida quyidagi bog’lanish mavjud: $Y = \alpha + \beta x + \varepsilon$

bu yerda $\alpha+\beta x$ tasodifiy bo’lmagan qismi, x erkli o’zgaruvchi sifatida qatnashadi, α va β lar esa aniqlanishi kerak bo’lgan noma'lum parametrlardir, ε- tasodifiy miqdor.

Regressiya tahlilining asosiy masalasi α va β parametrlarning bahosini va kuzatish natijalari bo’yicha o’tadigan to’g’ri chiziqning holatini aniqlashdan iboratdir.

Parametrlarni baholashning eng kichik kvadratlar usuli

Aytaylik X va Y lar uchun 4 ta kuzatish natijalariga egamiz va bu natijalar orqali α, β parametrlarning qiymatini aniqlaymiz.



Grafikda to’g’ri chiziqning Y o’qi bilan kesishish nuqtasi α ning bahosini bildiradi va *a* bilan belgilangan, to’g’ri chiziqning burchaк koeffitsienti esa β ning bahosini anglatib, *b* bilan belgilanadi. Birinchi qadam har bir kuzatishning xatosini aniqlashdan iboratdir.

$$\varepsilon_1 = y_1 - \hat{y}_1 \quad \varepsilon_2 = y_2 - \hat{y}_2 \quad \varepsilon_3 = y_3 - \hat{y}_3 \quad \varepsilon_4 = y_4 - \hat{y}_4$$

Quyilgan masalani yechishning usullaridan biri xatolar kvadratlarining yigindisini minimallashtirishdan iboratdir.

$$S = \varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2 + \varepsilon_4^2 \to \min$$

$$S(a,b) = \sum \varepsilon_i^2 \to \min$$

bundan quyidagini yozishimiz mumкin:

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

$$\begin{cases}\dfrac{\partial S}{\partial a}=-2\sum\limits_{i=1}^{n}(y_i-a-bx_i)=0\\[2ex]\dfrac{\partial S}{\partial b}=-2\sum\limits_{i=1}^{n}(y_i-a-bx_i)x_i=0\end{cases}$$

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

$$b = \frac{cov(x,y)}{var(x)}$$

$$a = \bar{y} - b\bar{x}$$

bu yerda

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

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

Misol. Regressiya tenglamasi noma'lum parametrlarini aniqlash va yechimini tahlil qilish

Regressiya tenglamasi parametrlarini bog'lanish turlariga qarab, eng kichik kvadratlar usuli bilan, normal tenglamalar sistemasini yechish orqali topiladi.

Bu usulni yuqorida keltirilgan tenglamalarga qo'llanilishini ko'rib chiqamiz:

1. Bir o'zgaruvchili bog'lanishda ya'ni, $y = (a_0 + a_1 x)$ da sistemaning ko'rinishi:

$$\begin{cases} na_o + a_1 \sum_i x_i = \sum_i y_i \\ a_o \sum_i x_i + a_1 \sum_i x_i^2 = \sum_i x_i \cdot y_i \end{cases}$$

Bu usuldan foydalanib, noma'lum parametrlarni aniqlashni quyidagi misol asosida ko'rib chiqamiz.

Misol. Quyida biz ihota daraxtzorlarining maydoni bilan tuproq erroziyasiga moyil bo'lgan tepaliklar qiyaliklarining tikligi orasidagi bog'lanishni ifodalaydigan to'g'ri chiziqli tenglamaning parametrlarini baholashni ko'rib chiqamiz. Regressiya tenglamasini quyidagi ko'rinishda qidiramiz.

$$y = (a_0 + a_1 x)$$

Qiyaliklar tikligi, grad. x	Oralik o'rtacha qiymati grad. x	Ihota daraxtzorlari maydoni, % y	x^2	X.Y	Y ning hisoblangan qiymati $\hat{y}$
1-2	1.5	2.85	2.25	4.35	2.71
2-2.5	2.25	2.98	5.06	6.75	3.32
2.5-3	2.75	3.65	7. 56	10.18	3.73
3-3.5	3.25	4.30	11.56	13.98	4.14
	9.75	13.78	25.43	35.26	—

Bu xolda tenglamalar sistemasining ko'rinishi

$$\begin{cases} a_o + 2{,}44a_1 = 3{,}45, \\ 2{,}44a_o + 6{,}36a_1 = 8{,}81 \end{cases}$$

dan iborat bo'ladi. Bu sistemani yechib, $a_0 = 1{,}48;$ $a_1 = 0{,}82$ arni topamiz.

U holda bog'lanish tenglamasining ko'rinishi bo'ladi $y = 1{,}48 + 0{,}82x$ bu yerda Y-ihota daraxtzorlari maydoni, %, X-haydaladigan yerlar qiyaligining o'rtachasi, grad.

Bu tenglama tahlili quyidagicha: agar haydaladigan yerlar qiyaligi 1 gradusga o'zgarsa ihota daraxtzorlari maydonini 0,82 % ga o'zgartirish kerak ekan.

Regresssiya tahlilning boshqa muhim masalasi, biz tanlagan model tanlama modelga teskari emas, ya'ni undan ko'p chetga chiqmasligi masalasini aniqlashdan iboratdir. Bunday masalaga modelning adekvatligini (tanlagan obyektga yaqinligini) tekshirish masalasi deyiladi. Matematik statistikada bu masalani yechish uchun juda ko'p usullar mavjud. Chiziqli regressiya modelining adekvantligini tekshiruvchi oddiy usuldan biri determinatsiya koeffitsientini tekshirishdir:

$$D = R^2$$

R2 1 ga qancha yaqin bo'lsa regressiya tenglamasining adekvatlik darajasi shuncha yuqori bo'ladi. Lekin R2 ning bitta kamchiligi shundaki, koeffitsientning katta qiymatlariga kuzatishlar soni kam bo'lgan hollarda erishiladi.

Regressiya koeffitsientlarining aniqligi

Umumiy holda aytish mumkinki, regressiya koeffitsiyenti, yuqori aniqlikka erishishi uchun quyidagilar bajarilishi kerak:

1. tanlamadagi kuzatishlar soni qancha ko’p bo’lsa;
2. tushuntiruvchi o’zgaruvchilarda tanlama dispersiyasi qancha katta bo’lsa;
3. tasodifiy hadning nazariy dispersiyasi qancha kam bo’lsa;

Regressiya koeffitsientlarining siljimasligi

Matematik statistikada parametrlarni baholash sifati α va β miqdorlarning siljimaslik miqdori bilan xarakterlanadi va u $M(a) = a$, $M(b) = b$ bo‘ladi.

Bu erda M(ξ) ξ-tasodifiy miqdorning matematik kutishi.

a va *b* ning asoslanganligi:

$$\lim_{n\to\infty} var(a) = 0, \quad \lim_{n\to\infty} var(b) = 0.$$

Bu baholashlarning sifati, ular qaysi usul bilan hosil qilinganligiga bog’liq. Bu yerda *a* va *b* baholarni hosil qilish uchun eng kichik kvadratlar usulidan foydalanildi. Matematik statistika kursida eng kichik kvadratlar usuli asosida olingan baholar siljimagan va asosli baholar deyiladi. Demak *a* va *b* lar siljimagan va asosli baholardir.

ETIBORINGIZ UCHUN RAHMAT!