

РЕГРЕССИЯЛЫҚ ТАЛДАУДЫҢ НЕГІЗГІ
ӘДІСТЕРІН ҚОЛДАНЫП
БИОЛОГИЯЛЫҚ ЖӘНЕ
МЕДИЦИНАЛЫҚ МАЗМҰНЫ

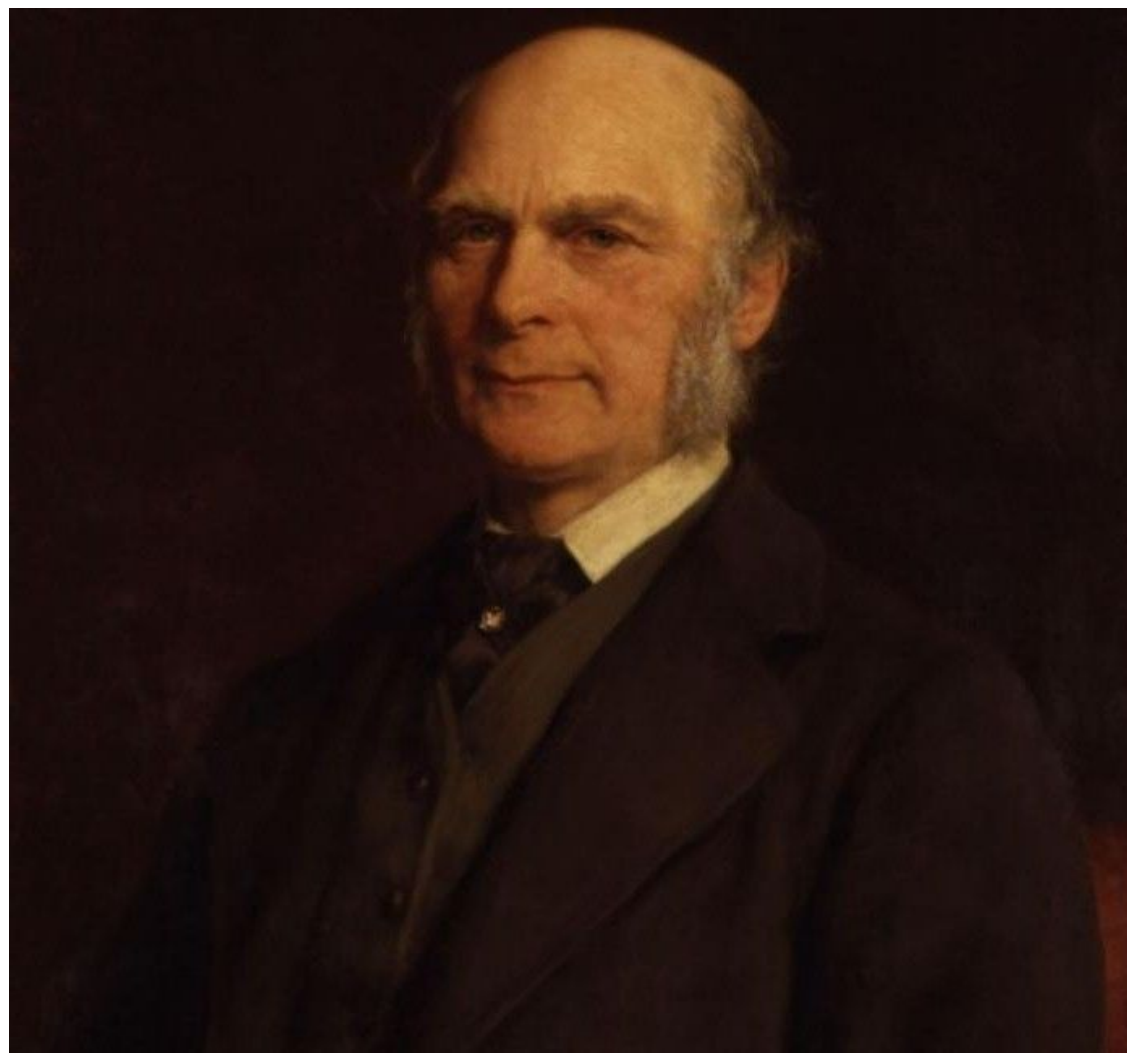
Жоспар:

- I Кіріспе
- II Негізгі бөлім
 - 2.1 Регрессия түсінігі
 - 2.2 Жұпталған қарапайым регрессия
 - 2.3 Көптік регрессия
 - 2.4 Регрессия теңдеуін құру
- III Қорытынды
- IV Пайдаланылған әдебиеттер

Кіріспе

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

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



Френсис Гальтон (1822-1911)

Регрессия түрлері

Белгілердің санына қарай регрессияны екіге бөледі. Олар:

Жұпталған
(қарапайым)

Көптік

Жұпталған регрессия

Жұпталған регрессия- екі факторлар арасында құрылатын модель. Мысалы, моделді құру кезінде тауардың тұтынымы кіріске байланысты өзгеретінін ескермесе, онда оны барлық факторларға бірдей әсері бар деп тұжырымдайды. Бірақ , кірістің тұтынуға әсерінің қаншалықты дұрыстығын анықтау үшін, басқа фактордың өзгеруінің корреляциясын анықтау қажет. Сондықтан, бұл жағдайда модельге енетін басқа да факторларды табу керек. Бұл жағдайда көптік регрессия теңдеуі құрылады

$y=f(x)$ регрессиялық талдау келесі кезеңдерден тұрады:

- Функция түрін анықтау;
- Регрессия коэффициенттерін анықтау;
- Нәтижелі белгінің теориялық мәндерін есептеу;
- Регрессия коэффициентінің статистикалық маңыздылығын тексеру;
- Регрессия теңдеуінің статистикалық маңыздылығын тексеру.

Көптік регрессия

- Көптік регрессия- теңдеуге басқа да факторлардың әсері болуы жағдайында құрылатын теңдеу
- Көптік регрессия сұраныс, кіріс, өндіріс алымдары функциясын, макроэкономикалық есептеулер мен басқа да эконометрикалық сұрақтарды шешу үшін қолданылатын модель. Эконометрикадағы негізгі әдіс - көптік регрессия әдісі. Көптік регрессия әдісін құру үшін алдымен моделдің құрылымын білу керек.

$$y=f(x_1,x_2,x_3, \\ \dots x_n)$$

Көптік регрессияны құру үшін алдымен оның моделінің құрылымын анықтау керек.

Ол 2 жағадайда болады:

- -факторларды таңдау;

- регрессия теңдеуінің түрін анықтау

Көптік регрессия әдісіне енетін факторлар мына талаптарға байланысты :

- Олар сандық жағынан өлшемді. Егер модельге сандық өлшемі болмайтын сапалы фактор енгізсек, онда оған сан жағынан анықтама беру қажет. Мысалы, тұқым сапасы ретінде балл, қозғалмайтын объект ретінде оның орны: аудан т,б, ескеріледі.
- Факторлар тура функциональді байланыста болуы керек.
- Көптік регрессияға енетін факторлар айнымалының вариациясын білдіреді.

- Егер p факторы бар модель құрылса, онда ол үшін детерминация көрсеткіші R^2 детерминация көрсеткіші есептелуі тиіс, яғни шешуші белгінің вариациясын анықтау керек. Модельге кірмейтін факторларды $1-R^2$ түріндегі қалдық дисперсияға S_2 байланысты шешу керек. Егер регрессияға қосымша $p+1$ факторлары енсе, онда детерминация коэффициенті өседі, ал қалдық дисперсия кемиді, яғни
- Егер бұл шарт орындалмаса, онда енгізілген факторлар модельдің дұрыс құрылғанын көрсетпейді, сондықтан бұл факторлар артық деп саналады.

Артық факторлары бар модель қалдық дисперсияның шамасын кемітпейді және t-Стьюдент критерийі бойынша статистикалық мәнсіздікке әкеліп соғады. Сондықтан, факторларды таңдау теориялық-экономикалық талдаудың сапалылығына байланысты болады. Факторларды таңдау екі стадияда өтеді:

- **-негізгі факторлар**
- **-корреляция көрсеткішінің матрицасы**

Байланысты сипаттау үшін келесі жұпталған регрессия теңдеулерінің түрлерін қолданады:

- $y = a + bx$ – сызықты;
- $y = e^{ax + b}$ – экспоненциалды;
- $y = a + b/x$ – гиперболалы;
- $y = a + b_1x + b_2x^2$ – параболаалы;
- $y = ab^x$ – көрсеткіштік және т.б.

Регрессия теңдпуін құру оның коэффициенттерін (параметрлерін) бағалауға әкеліп соқтырады, ол үшін ең кіші квадраттар әдісін қолданады (ЕКӘ)

- Ең кіші квадраттар әдісі нақты нәтижелік белгі у мәнінің ух теориялық мәнінен ауытқу кадратының қосндысы ең аз болатын параметрлерді бағалауға мүмкіндік береді, яғни

$$\Sigma (y - y_x)^2 \rightarrow \min .$$

Ең кіші квадраттар әдісі бойынша $y=a+bx$

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

- мұндағы a – еркін коэффициент, b – регрессия коэффициенті, бірлік өлшемде факторлық белгі x өзгергенде, нәтижелі белгі y қаншаға өзгертіндігін көрсетеді

Регрессия коэффициентінің статистикалық маңыздылығын бағалау үшін Стьюденттің t-белгісі қолданылады

Регрессия коэффициентінің маңыздылығын тексерудің сызбасы:

1) H_0 : $a = 0, b = 0$ - регрессия коэффициенттері нөлден елеусіз өзгеше.

H_1 : $a \neq 0, b \neq 0$ - регрессия коэффициенттері нөлден елеулі өзгеше.

2) $p = 0,05$ - маңыздылық деңгейі.

$$3) t_{b \text{ есен}} = \frac{b}{m_b}, \quad t_{a \text{ есен}} = \frac{a}{m_a},$$

мұндағы m_b, m_a - кездейсоқ қателіктер.

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

4) $t_{кесте}(p; f)$,

5 кесте

мұндағы $f = n - k - 1$ - еркіндік дәрежесінің саны, n – бақылау саны, k – теңдеудегі « x » айнымалы болғандағы параметрлер саны.

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

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

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

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

Қорытынды