ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
БІЛІМ ЖӘНЕ ҒЫЛЫМ МИНИСТРЛІГІ

Қ. И. Сәтбаев атындағы
Қазақ ұлттық техникалық зерттеу университеті

Р. Т. Исмаилова

ЛОГИСТИКАЛЫҚ ЖҮЙЕЛЕРДЕГІ БОЛЖАУ ӘДІСТЕРІ

Қазақстан Республикасы Білім және ғылым министрлігі
оқулық ретінде ұсынған

Алматы 2017

ӘОЖ 656.(075)
КБЖ 39.18я73
И 88

Пікір жазғандар:
Ж. Х. Давильбекова, экон. ғыл. док., Қ. И. Сәтбаев атындағы ҚазҰТЗУ профессоры;
И. У. Утепбергенов, техн. ғыл. док., Алматы энергетикалық университетінің профессоры;
Б. Б. Мананов, экон. ғыл. канд., Т. Рысқұлов атындағы Жаңа экономикалық университеті доценті;
А. Т. Ускеленова, экон. ғыл. док., PhD, Білімді басқару зертханасының ғылыми қызметкері, Алматы Менеджмент Университеті.

Қазақстан Республикасы Білім және ғылым министрлігінің 2017 жылғы басылым жоспары бойынша басылды

Исмаилова Р. Т.
И 88 Логистикалық жүйелердегі болжау әдістері Алматы: Оқулық / *Р. Т. Исмаилова*, ҚазҰТЗУ, 2017. – 181 б. Сурет 24. Кесте 77. Әдебиеттер тізімі – 39 атау.
ISBN 978-601-323-059-7

Оқулықта жұптық және жиынтық регрессиялар (ең кіші квадраттар әдісі, сызықтық байланысты бағалау, корреляция және детерминация коэффициенттері, коэффициенттердің статистикалық мәнділігін талдау, сенімділік интервалдары және сызықтық регрессиялық талдауда гипотезаларды тексеру), уақыттық қатарлар сияқты негізгі бөлімдер қарастырылған. Сондай-ақ логистикалық есептерді болжау және бір өлшемді уақыттық қатарларды талдаудың статистикалық әдістері қарастырылған.

Кітап экономикалық мамандықтардың студенттеріне, сондай-ақ логистикалық жүйелердегі болжау әдістерін зерттейтін магистрантарға, оқытушыларға арналған.

ӘОЖ 339.18.(075)

ISBN 978-601-323-059-7

© Исмаилова Р. Т., 2017
© ҚазҰТЗУ, 2017

КІРІСПЕ

Болжау жасаудың статистикалық әдістері тек экономикада ғана емес, логистикада да маңызды орын алды. Болжау жасаудың статистикалық әдістері қамсыздандыру, өндірістік, тарату логистикасында бүгінде кең қолданыла бастады.

Қазіргі кезде көлемді есептеулерді, кестелер мен графиктерді қолмен жасау керек емес – бұл қара жұмыстың барлығын компьютер орындайды. Адамға зерттеу және шығармашылық жұмыстар қалады, олар: есептің қойылымы, алынған модельдердің сапасын бағалау, нәтижелердің интерпретациясы. Қазіргі кезде ақпараттың үлкен көлемін өңдеудегі компьютерлердің мүмкіндігі қолданушылар үшін статистикалық талдаудың қазіргі әдістерін қолжетімді етті. Қазіргі кезде бағдарламалардың статистикалық және эконометрикалық пакеттері құрастырылды, олар: STATISTICA, SPSS, SAS, STATGRAPHICS PLUS, МЕЗОЗАВР, STADIA, EViews бағдарламалық жүйесі, Novo Forecast, сондай-ақ себеп-салдар байланыстарын түсіндіретін есептерге, әлеуметтік болжау есептеріне, үрдістерді оңтайландыру есептеріне қолданыла алатын – MS Excel.

Оқу материалын жақсы түсіну үшін математика, ықтималдықтар теориясы және математикалық статистика, статистика, эконометрика, логистика негіздері және компьютерлік технология сияқты салалардан білім қажет.

Берілген оқулықта жүйелендірілген түрде бір өлшемді уақыттық қатарларды және логистикалық есептерге болжау жасауды талдаудың статистикалық әдістері берілген.

1. ЭКОНОМЕТРИКАҒА КІРІСПЕ

Эконометрика – экономикадағы математикалық статистиканың әдістерімен сандық заңдылықтарын және өзара тәуелділіктерін зерттейтін ғылым. Эконометрика экономикалық теориямен дәлелденген экономикалық заңдылықтарды статистикалық және математикалық әдістер көмегімен талдайды. Эконометрикалық талдауға негізделген экономикалық шешімдерді қабылдау – экономикалық талдау мен болжау үшін негіз болып табылады. Эконометриканың мақсаты – экономикалық заңдардың эмпирикалық нәтижесін құрастыру. Эконометриканың міндеті – экономикалық модельдер құру және олардың параметрлерін бағалау, экономикалық көрсеткіштердің қасиеттері туралы гипотезаларды және олардың байланыстары формаларын тексеру [1,3,4,5,9].

Барлық жағдайда модельдерді қолданудың маңызы жоғары. Талдау үшін және болжау үшін қолданатын модельдердің үш түрі бар: уақыттық қатарлар; бір теңдеулі регрессиялық модельдер; бір уақыттық теңдеулер жүйесі. Егер модельде тек бір ғана түсіндіруші айнымалы болса, яғни k =1, онда ол *жұп регрессия* деп аталады, k>1 болса, *жиынт*ық *регрессия* деп аталады.

Эконометриялық модельдеудің негізін статистикалық деректер құрайды. Олар типі бойынша бөлінеді. *Қиылыстық деректер* уақытқа тәуелді болмаған жағдайда, бір уақытта немесе әртүрлі уақытта алынған әртүрлі объектілер үшін деректер (фирма, аймақтар, шаруашылықтар) экономикалық көрсеткіштер бойынша жиналады. *Уақыттық қатарлар* бір объект үшін әртүрлі уақытта алынған деректер. Осы екі аралықтағы алынған деректер *панельді деректер* деп аталады, олар аса үлкен емес объектілер үшін алынған аз уақыттағы деректер, мысалы Қазақстандағы кәсіпорындардың соңғы үш жылдағы табыстары болуы мүмкін.

1.1. Ықтималдық теориясының және математикалық статистиканың элементтері

Кездейсоқ айнымалы деп әрбір берілген жиыннан нақты ықтималдықпен мән қабылдай алатын айнымалыны айтамыз. Кездейсоқ шамалар дискретті немесе үздіксіз болып бөлінеді. Егер кездейсоқ шама ақырғы және санаулы сандық мәндерді қабылдаса, онда бұл *дискретті* шама. Мысалы, екі ойын сүйегін лақтырғандағы түскен ұпайлар саны; дүкенде бір күннің ішінде сатылған телевизорлар саны [1,3,9,13,14].

Бұлардан келесідей кесте түрінде көрсетуге болады:

x	x_1	x_2	….	x_n
p	p_1	p_2	….	p_n

Мұндағы x_1, x_2, … x_n – x шаманың барлық мәндері p_1 $p_2 \ldots p_n$ ықтималдықтарына тең.

Және бұл: $\sum_{i=1}^{n} p_i = 1.$

Қандайда бір интервалда кез келген мән қабылдайтын кездейсоқ шама – *үздіксіз* кездейсоқ шама болып табылады. Мысалы, бөлмедегі температура; тәулік бойындағы сауда орнындағы доллардың биржалық бағасы ықтималдық тығыздығының функциясымен беріледі.

X кездейсоқ шамасының [a,b] интервалында мәндер қабылдау ықтималдығы төмендегі формулаға тең:

$$P\{a \le x \le b\} = \int_a^b p(x)dx.$$

Бұл [a,b] кесіндісіндегі P(x) ықтималдығының тығыздығы қисық ауданына тең. X-тің қандай да бір мәні іске асырылғанда

$$P\{-\infty < x < \infty\} = \int_{-\infty}^{\infty} p(x)dx = 1.$$

Логистикада ықтималдықтарды таратудың келесі заңдары кең тараған болып табылады: нормальді, экспоненциалды, биноминалды, Пуассон [13,14,32].

Биномиалды заң деп әрқайсында оқиғаның пайда болу ықтималдығы p-ға тең *n* тәуелсіз тәжірибелерде *x* оқиғалардың пайда болу саны – дискретті кездейсоқ шама болатын тарату заңы аталады; мұнда *x=k* (оқиғалардың пайда болу саны) мүмкін мәндер ықтималдығы Бернулли формуласы бойынша есептеледі: $P_n(k) = C_n^k p^k (1-p)^{n-k}$, мұнда *n* – натуралды сан, k=0,1,2,…,n.

Келесі белгілеу қабылданған $1-p=q$, $P_n(k) = C_n^k p^k q^{n-k}$

$$C_n^k = \frac{n!}{k!(n-k)!} = \frac{n(n-1)(n-2)...(n-(k-1))}{k!}.$$

1.1-мысал: База 10 дүкенге қызмет көрсетеді. Бір дүкеннен тапсырыс келу ықтималдығы – 0,4. Тапсырыстың ең ықтималды санын, тапсырыстардың болуының ең ықтималды санын, 5 дүкеннен келетін тапсырыстардың келу ықтималдығын анықтау.

Есептеулер:

1) $N = np = 10 * 0{,}4 = 4$.

2) Тапсырыстың ең ықтималды саны Бернулли формуласы бойынша анықталады:

$$P_{n,m} = C_m^n p^m \cdot q^{mn-m}$$

$$P_{10,4} = C_{10}^4 \cdot 0{,}4^4 \cdot 0{,}6^6$$

$$C_{10}^4 = \frac{10!}{4! \cdot 6!} = \frac{7 \cdot 8 \cdot 9 \cdot 10}{1 \cdot 2 \cdot 3 \cdot 4} = 210$$

$$P_{10,4} = 210 \cdot 0{,}4^4 \cdot 0{,}6^6 = 0{,}251$$

3) 5 дүкеннен келетін тапсырыстардың келу ықтималдығы:

$$P_{10,5} = C_{10}^5 \cdot 0{,}4^5 \cdot 0{,}6^5$$

$$C_{10}^5 = \frac{10!}{5! \cdot 5!} = \frac{6 \cdot 7 \cdot 8 \cdot 9 \cdot 10}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} = 252$$

$$P_{10,5} = 252 \cdot 0{,}4^5 \cdot 0{,}6^5 = 0{,}201$$

Ықтималдықтардың тарату графигін тұрғызамыз, ол үшін дүкендерден түсетін тапсырыстардың қалған ықтималдықтарын

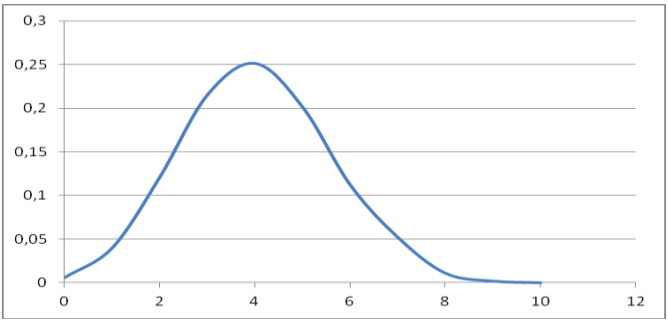
есептеу керек:

$$P_{10,0} = C_{10}^{0} \cdot 0{,}4^{0} \cdot 0{,}6^{10} = 0{,}006$$
$$P_{10,1} = C_{10}^{1} \cdot 0{,}4^{1} \cdot 0{,}6^{9} = 0{,}04$$
$$P_{10,2} = C_{10}^{2} \cdot 0{,}4^{2} \cdot 0{,}6^{8} = 0{,}121$$
$$P_{10,3} = C_{10}^{3} \cdot 0{,}4^{3} \cdot 0{,}6^{7} = 0{,}215$$
$$P_{10,6} = C_{10}^{6} \cdot 0{,}4^{6} \cdot 0{,}6^{4} = 0{,}112$$
$$P_{10,7} = C_{10}^{7} \cdot 0{,}4^{7} \cdot 0{,}6^{3} = 0{,}121$$
$$P_{10,8} = C_{10}^{8} \cdot 0{,}4^{8} \cdot 0{,}6^{2} = 0{,}011$$
$$P_{10,9} = C_{10}^{9} \cdot 0{,}4^{9} \cdot 0{,}6^{1} = 0{,}002$$
$$P_{10,10} = C_{10}^{10} \cdot 0{,}4^{10} \cdot 0{,}6^{0} = 0{,}0001$$

Мәліметтерді 1.1-кестеге енгіземіз, есептелген мәндер негізінде график тұрғызамыз (1.1-сурет):

1.1-кесте

M	0	1	2	3	4	5	6	7	8	9	10
P(10,m)	0,006	0,04	0,121	0,215	0,251	0,201	0,112	0,121	0,011	0,002	0,0001



1.1-сурет. Ықтималдықтардың тарату графигі

Есептеулер жасалып, тапсырыстардың ең ықтималды саны анықталды $N = 4$. Бұл базаға 4 тапсырыс келетінін білдіреді. Тапсырыс келуінің ең ықтималды саны $P_{10,4} = 0{,}251$. Базаға тапсырыстардың түсуінің есептелген ықтималдықтарынан $P_{10,4} = 0{,}251$ ықтималдығы басқаларға қарағанда ең үлкен

болады, мұны ықтималдықтарды тарату графигінен көруге де болады. $m = 4$ дейін базаға тапсырыстардың келу ықтималдығы өседі, кейін кемиді. 5 дүкеннен тапсырыстардың келу ықтималдығы $P_{10,5} = 0{,}201$ және $P_{10,5} < P_{10,4}$.

1.1-тапсырма. Жеткізілім партиясында 100 бұйым бар, оның 5 % (k = 5) кемістігі бар бұйымдар. 20 бұйымнан тұратын іріктемеден кемістігі бар бұйымның анықталу ықтималдығын анықтау, яғни[32]:

N	*k*	*M*	*L*	*P*
100	5	20	1	?

Бірінші және екінші жағдайда алынған ықтималдықтарды түсіндіру. Кемістігі бар бұйымды анықтау ықтималдығы іріктемені 4 есе 5-тен 20-ға артқан кезде қалай өзгеретінін байқау.

1.2-тапсырма. База 10 тұтынушыны қамтамасыз етеді. Бір тұтынушыдан тапсырыс түсу ықтималдығы 0,7-ге тең. 0, 1, 2, 3…10 тапсырыстардың келу ықтималдығын анықтау.

1.3-тапсырма. Тарату орталығы 20 дүкенге қызмет көрсетеді. Тауарға тапсырыс келу ықтималдығы 0,6.

Анықтау керек:

– тапсырыстың ең ықтималды саны;

– тапсырыстың ең ықтималды санының түсу ықтималдығы; 19

1.4-тапсырма. Анықтау керек:

– тапсырыстың ең ықтималды саны;

– тапсырыстың ең ықтималды санының түсу ықтималдығы; 19

Берілгені: Өнімді жүзеге асыру мәліметтері (1.2-кесте)

1.2-кесте

Өнімді жүзеге асыру мәліметтері

Жүзеге асыру көлемі, млн. тг.	1	1–2	2–3	3–4	4–5	5–6	6 артық
Жағдайлар саны	2	4	6	10	8	3	2

Анықтау керек:
– параметрлері;
– ықтималдықты тарату заңының графигін тұрғызу.

Бақылау сұрақтары

1. Эконометрикаға анықтама беріңіз.
2. Кездейсоқ шамаға анықтама беріңіз.
3. Статистикалық мәліметтердің негізгі түрлерін атаңыз.
4. Дискретті кездейсоқ шамаға анықтама беріңіз.
5. Үздіксіз кездейсоқ шамаға анықтама беріңіз.

1.2. Жиынтық сандық сипаттамалары

Кездейсоқ *x* шаманың жалпы жиындағы сандық сипаттамалары үшін оның математикалық үміті және теориялық дисперсиясын қарастырамыз [1,3,9].

Дискретті кездейсоқ шаманың математикалық үміті оның барлық мүмкін болатын мәндерін бастапқы сәйкес ықтималдықпен салыстыру. Егер кездейсоқ шаманы математикалық түрде x деп белгілесек, оның математикалық үміті E(x) деп белгіленеді.

X шамасы n нақты мәндерін ($x_1, x_2, \ldots, x_n$) қабылдаса x_i-ді алу ықтималдығы p_i-ға тең. Онда дискретті кездейсоқ шаманың математикалық үміті келесіге тең:

$$E(x) = x_1 p_1 + x_2 p_2 + \ldots + x_n p_n = \sum_{i=1}^{n} x_i p_i \qquad (1.1)$$

мысалы, үлестіруі үшін математикалық үміті: E(x)=1*0,3+3*0,5+5*0,2=2,8.

x	1	3	5
p	0,3	0,5	0,2

Кездейсоқ шаманың математикалық үміті оның жалпы жиынтықтығының орташа шамасы деп аталады. Кездейсоқ

айнымалының барлық мүмкін болатын мәндерінің жиынтығы жалпы жиын деп аталады.

Х-кездейсоқ шамасы үшін бұл мән μ деп белгіленеді.

1-қасиет. Кездейсоқ шама қосындысының математикалық үміті олардың математикалық үміттерінің қосындысына тең, яғни,

$$E(x+y+z)=E(x)+E(y)+E(z) \qquad (1.2)$$

2-қасиет. Егер кездейсоқ айнымалы тұрақтыға көбейсе, онда оның математикалық үміті дәл сол тұрақтыға көбейеді. Егер х-кездейсоқ айнымалы, ал *а*-тұрақты болса, онда

$$E(ax)=aE(x) \qquad (1.3)$$

3-қасиет. Тұрақтының математикалық үміті дегеніміз – $E(a)=a$ Х-кездейсоқ шамасының (теориялық) дисперсиясы деп – *х*-ң $\mu=E(x)$, математикалық үмітінің ауытқуын атаймыз, яғни,

$$\sigma_x{}^2=E\{(x-\mu)^2\} \qquad (1.4)$$

дискретті кездейсоқ шама үшін:

$$\sigma_x{}^2=\sum_{i=1}^{n}(x_i-\mu)^2 p_i,$$

үздіксіз кездейсоқ шама үшін

$$\sigma_x{}^2=\int_{-\infty}^{\infty}(x-\mu)^2 p(x)dx.$$

Дискретті кездейсоқ шама үшін дисперсияны есептеу

$\sigma_{xy}{}^2=(1-2{,}8)^2*0{,}3+(3-2{,}8)^2*0{,}5+(5-2{,}8)^2*0{,}2=1{,}96$,

$\sigma_{xy}=1{,}4$.

σ_{xy}-теориялық стандарты ауытқу – сол ауытқудан квадраттық түбірді алу арқылы есептелінеді. σ_{xy}-ң мәні оң болса, онда х, ү арасында тура статистикалық байланыстың барлығын дәлелдейді, ал σ_{xy}-ң теріс мәнге ие болса, х, ү арасында кері статистикалық байланыстың бар екенін көрсетеді.

1-қасиет. Тұрақтының дисперсиясы нөлге тең, яғни $\sigma^2(c)=0$, мұнда c – константа.

2-қасиет. Кездейсоқ шаманы константаға көбейткен кезде оның дисперсиясы бұл константаның квадратына көбейтіледі $\sigma^2(cx)=c^2\sigma^2(x)$.

3-қасиет. Тәуелсіз кездейсоқ шамалардың сомаларының дисперсиясы олардың дисперсияларының сомасына тең

$$\sigma^2(x+y+z)=\sigma^2(x)+\sigma^2(y)+\sigma^2(z).$$

4-қасиет. $\sigma^2(a+bx)=b^2\sigma^2(x)$, мұнда a,b – тұрақтылар.

5-қасиет. Биномиальді таратудың дисперсиясы $\sigma^2(x)=npq$, мұндағы n –тәуелсіз тәжірибелер саны;

p – бір тәжірибедегі оқиғаның пайда болу саны;

$q=1-p$ бір тәжірибедегі оқиғаның пайда болмау саны.

Кездейсоқ шаманың нормальді таратылымы екі параметрмен сипатталады: μ орташа мәнімен және σ^2 дисперсиямен.

Бұл былай белгіленеді: $x\sim(\mu,\sigma^2)$.

Математикалық үміті нөлге тең, ал дисперсиясы бірге тең болатын нормальді таратылым стандартты нормальді үлестірілім деп аталады және $x\sim(\mu,\sigma^2)$ түрінде жазылады.

1.5-тапсырма. *x* кездейсоқ шамасы екі ойын сүйегін лақтырғанда түскен үлкен және кіші екі санның айырмасы ретінде анықталады. Егер олар бір-бірімен тең болса, онда x айнымалысы нөлге тең. *x* үшін ықтималдықты үлестіруді және математикалық үмітті тауып $E(x^2)$ есептеңіз [10].

1.6-тапсырма. *x* – μ-математикалық үміті бар кездейсоқ шама, ал λ –тұрақты. *x* математикалық үміті λμ-ге тең екенін дәлелдеңіз.

1.7-тапсырма. x – ойын сүйегін бір рет лақтырғандағы ұпайлар саны болсын. *y*-ті анықтаңыз, мұндағы y келесі $y=x^2+3x-2$ формула бойынша есептелінеді және одан кейін

Е(у) шамасын есептеңіз, ол $E(x^2)+3E(x)-2$-ке теңелетінін дәлелдеңіз.

1.8-тапсырма. Теориялық дисперсияны және *x*-тің стандартты ауытқуын есептеңіз (1.5-тапсырмадағыдай).

Бақылау сұрақтары

1. Математикалық үмітке анықтама беріңіз, оның қасиеттерін келтіріңіз.
2. Теориялық стандарттық және орташа квадраттық ауытқуға анықтама беріңіз.
3. Басты жиынтыққа анықтама беріңіз.

1.3. Кездейсоқ айнымалының тұрақты және айнымалы құрамдастары

Көп жағдайда кездейсоқ шаманы бір бүтін ретінде қарастырудың орнына оны тұрақты және таза кездейсоқ құрамдастарға бөлген ыңғайлы, мұнда тұрақты құрамдас оның математикалық үміті. Егер x – кездейсоқ айнымалы болса және μ – оның математикалық үміті, онда кездейсоқ шаманың декомпозициясы келесі түрде жазылады: $x = \mu + u$, мұнда u – таза кездейсоқ құрамдас [1,3,9].

Кездейсоқ құрамдас x және μ аралығындағы айырым ретінде анықталады: $u = x - u$. Бұл анықтамадан u шамасының математикалық үміті 0-ге тең, ал x-ң теориялық дисперсиясы u-ң теориялық дисперсиясына тең екенін көруге болады.

Бағалау әдістері және бағалау

Осы уақытқа дейін кездейсоқ айнымалы туралы нақты ақпарат бар деп қарастырдық, дәлірек айтқанда оның ықтималдық үлестірілуі немесе үлестіру тығыздығының функциясы туралы ақпарат көп тараған.

Бұл ақпарат көмегімен теориялық математикалық үмітті, дисперсияны және кез келген басқа сипаттамаларды есептеуге болады.

Алайда, іс жүзінде жасанды кездейсоқ шамаларды ескермегенде біз ықтималдықты, үлестіруді және ықтималдықты үлестіру тығыздығын нақты түрде біле алмаймыз. Бұл дегеніміз математикалық күту және дисперсия белгісіз дегенді білдіреді. Ал бізге басты жиынның осы және басқа сипаттамаларын бағалауымыз қажет болып жатады.

Бағалау процедурасы әрқашан бірдей: n байқау жүргізіледі және сәйкес формула бойынша қажетті сипаттаманың бағасы есептелінеді. Бағалау әдісі – ортақ ереже немесе формула, ал бағалау мәні – таңдаудан келесі таңдауға дейін өзгеріп отыратын нақты сан.

Басты жиынның негізгі екі сипаттамаларын бағалау формулаларын келтірейік.

Орташа таңдамалы : $\bar{x} = \frac{1}{n}\sum x_i$;

Кездейсоқ айнымалының таңдамалы дисперсиясы:

$$s^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2 .$$

1.2-мысал. 2000−2005 жылдарғы Қазақстанның физикалық көлемінің өсу қарқыны ЖШӨ келесі негізде анықталады: 9.8, 13.5 ,9.3,9.4,9.4.

Физикалық көлемінің орташа өсу қарқынын ЖШӨ табамыз:

$$\bar{x} = \frac{1}{6}(9{,}8 + 13{,}5 + 9{,}3 + 9{,}3 + 9{,}4 + 9{,}4) = 10{,}12.$$

Кездейсоқ айнымалының әйтеуір бір әдісімен алынған бағаның өзі де кездейсоқ айнымалы болып табылады, өйткені ол айнымалының кездейсоқ іске асуына негізделген.

Кездейсоқ айнымалының сипаттамасының бағасы *ығыстырылмаған* деп аталады, егер оның математикалық үміті бұл сипаттаманың теориялық мәніне сәйкес келсе.

Мысалы, $\mu = E(x)$ математикалық үміті үшін $\bar{x}$ орташа таңдамалы бағасы бола алады. Енді

$$E(\bar{x}) = E(\frac{1}{n}\sum_{i=1}^{n} x_i) = \frac{1}{n}\sum_{i=1}^{n} E(x_i) = \mu \ .$$

Яғни, $\bar{x}$ – х кездейсоқ шамасының μ математикалық үміті үшін ығыстырылмаған баға болып табылады.

$\lambda_1 + \lambda_2 = 1$ кезінде $z = \lambda_1 x_1 + \lambda_2 x_2$ жалпыланған бағасы μ математикалық үміті үшін ығыстырылмаған баға болады. Бұл $\lambda_1, \lambda_2, ...\lambda_n$ кез келген санына ақиқат.

Тиімді баға – дисперсиясы минималды болатын баға. Теориялық орташаның жалпыланған бағасының дисперсиясын қарастырайық, және ол екі зерттеудің де тең салмағы болған жағдайда минималды болатынын көрсетеміз.

Егер x_1 және x_2 зерттеулер тәуелсіз болса, жалпыланған бағасының теориялық дисперсиясы мынаған тең болады:

$$pop.var(Z) = pop.var(\lambda_1 x_1 + \lambda_2 x_2) = (\lambda_1^2 + \lambda_2^2)\sigma^2 \qquad (1.6)$$

Біз бағаның ығыстырылмағандығы үшін λ_1 және λ_2 сомасы бірге тең болуы қажет екенін айқындадық. Бұдан шығатыны, ығыстырылмаған бағалар үшін $\lambda_2 = (1 - \lambda_1)$ және

$$\lambda_1^2 + \lambda_2^2 = \lambda_1^2 + (1 - \lambda_1)^2 = 2\lambda_1^2 - 2\lambda_1 + 1 \qquad (1.7)$$

Біз дисперсияны минимизациялау үшін λ_1 таңдағымыз келетіндіктен, бізге $(2\lambda_1^2 - 2\lambda_1 + 1)$ минималдау керек. Минимум $\lambda_1 = 0.5$ кезінде болады. Бұдан шығатыны, λ_2=0,5.

Сонымен, біз қарастырылып отырған бағалардың ішінде таңдамалы орташаның дисперсиясы ең аз екенін көрсеттік.

Егер ықтималдық бойынша бағаның шегі басты жиынның сипаттамаларының ақиқат мәніне тең болса, онда бұл баға *тиянақты* болып табылады.

Іріктеменің өлшемі өскен кезде таратылым ақиқат мәнінің айналасына симметриялы болады, бұл асимптотикалық ығыстырмағандықты көрсетеді. Оның ақырғы есепте ақиқат

мәннің жалғыз нүктесіне айналатыны оның тиянақтылығы туралы айтады [3,9].

1.9-тапсырма. 1.3-кестеде 3 көрсеткіш бойынша таңдамалы мәліметтер берілген. Орташа таңдамалыны, дисперсияны, қалыпты ауытқуларды табыңыз [7,10].

1.3-кесте

i	1	2	3	4	5	6
x	9	6	5	4	5	3
y	2	3	6	6	7	9

1.10-тапсырма **(**$\lambda_1 x_1$+...+ **(**$\lambda_n x_n$**)** жалпыланған формуласы μ ығыспаған бағасын беретін n бақылау λ_1+...+ λ_n=*1* болып табылады.

1.11-тапсырма *x* кездейсоқ шамасы 3 және 4 мәндерін қабылдайды, дәл осындай ықтималдықпен *y* кездейсоқ шамасы 1 және 2 мәндерін қабылдайды x және y бір-бірінен тәуелсіз түрде үлестірілген (1.4-кесте). *Z* айнымалысы *Z=x/y* арқылы анықталады және *p* қайсысы 0,25 ықтималдыққа тең төрт мүмкін болатын мәнге ие:

1.4-кесте

X У	3 1	4 2
1 2	3,0 1,5	4,0 2,0

E(Z) шамасы E (x)/E(y) өрнегіне тең емес екенін көрсетіңіз.

Бақылау сұрақтары

1. Нақты u кездейсоқ құрамдас дегеніміз не?
2. Баға дегеніміз не?
3. Бағалау әдісі не үшін керек?
4. Қандай бағалау тиімді деп саналады?

5. Қандай бағалау ығыстырылмаған деп саналады?
6. Қандай бағалау тиянақты орындалды деп саналады?

1.4. Ковариация, дисперсия және корреляция коэффициенті

Ковариация екі айнымалы арасындағы өзара байланыс өлшемі болып табылады [3,4,6,9].

Ковариацияның *таңдамалы* және *теориялық* болып бөлінеді.

Егер x пен y – кездейсоқ айнымалылар болса, онда *теоретикалық ковариация* σ_{xy} бұл көрсеткіштердің ауытқуларының олардың орташа мәндерінінен айырмаларының көбейтіндісінің математикалық үміті ретінде анықталады:

$$\sigma_{xy} = E\{(x-\mu_x)(y-\mu_y)\},$$

мұндағы μ_x пен μ_y – х пен у теоретикалық орташа мәндері.

Ковариация келесі формуламен беріледі:

$$Cov(x,y) = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) = \frac{1}{n}\{(x_1-\bar{x})(y_i-\bar{y})+...+(x_n-\bar{x})(y_n-\bar{y})\}$$

немесе $\operatorname{cov}(x,y) = \overline{xy} - \bar{x}\cdot\bar{y}$

Ковариация есептеу ережелері:

1. Егер $y = v + w$, онда

$Cov(x,y) = Cov(x,v) + Cov(x,w)$.

2. Егер $y = az$, мұнда a – константа, онда

$Cov(x,y) = aCov(x,z)$.

3. Егер $y = a$, мұнда a – константа, онда $Cov(x,y) = 0$.

4. $\operatorname{cov}(x,x) = \frac{1}{n}\sum(x_k-\bar{x})^2 = \operatorname{var}(x)$

5. $\operatorname{cov}(x,y) = \operatorname{cov}(y,x)$

Егер x және y – кездейсоқ шамалар болса, онда σ_{xy} теориялық ковариация осы шаманың олардың орташа шамасынан ауытқуы көбейтіндісінің математикалық күтуі ретінде анықталады: $\sigma_{xy} = E\{(x-\mu_x)(y-\mu_y)\}$, мұнда μ_x и μ_y – x және y теориялық үміті.

Теоретикалық ковариацияны есептеу ережелері таңдамалы ковариацияны есептеу ережелеріне ұқсас.

Тандамалы дисперсия. *Таңдамалы дисперсия (немесе вариация)* дегеніміз кездейсоқ шаманың ауытқу квадраттарының арифметикалық орташасы:

$$S_x^2 = \mathrm{var}(x) = \frac{1}{n}\sum_{k=1}^{n}\left(x_k - \overline{x}\right)^2$$

Таңдамалы дисперсия келесі формуламен анықталады: $S_x^2 = \mathrm{var}(x) = \overline{x}^2 - \left(\overline{x}\right)^2$

Таңдамалы дисперссиялық есептеу ережелері:

1. Егер $y = v + w$, онда $Var(y) = Var(v) + Var(w) + 2Cov(v,w)$.

2. Егер $y = az$, *a*- тұрақты, онда $Var(y) = a^2 Var(z)$.

3. Егер $y = a$, *a*-тұрақты болса, онда $Var(y) = 0$.

4. Егер $y = v + a$, *a*- тұрақты.болса, онда $Var(y) = Var(v)$.

Іріктемедегі кездейсоқ шаманың стандартты немесе орташа квадраттық ауытқуы дегеніміз $S = \sqrt{S^2}$ көрсеткіші.

Таңдамалы дисперсия мен іріктемелік ковариация байланысы $y = v + w$ болғанда, $Var(y) = Var(v) + Var(w) + 2Cov(v,w)$ өрнегімен анықталады.

1.3-мысал. 2000−2005 жылдарғы Қазақстанның физикалық көлемінің өсу қарқыны ЖІӨ келесі негізде анықталады: 9.8, 13.5, 9.3,9.4,9.4.

Физикалық көлемінің орташа өсу қарқынын ЖІӨ табамыз:

$$\bar{x} = \frac{1}{6}(9{,}8 + 13{,}5 + 9{,}3 + 9{,}3 + 9{,}4 + 9{,}4) = 10{,}12.,$$

Таңдамалы вариация үшін төмендегідей:

$$Var(x) = \frac{1}{6}\left[\begin{array}{l}(9.8-10.12)^2 + (13.5-10.12)^2 + (9.3-10.12)^2 + (9.3-10.12)^2 + \\ + (9.4-10.12)^2 + (9.4-10.12)^2\end{array}\right] = 2.32$$

Таңдамалы дисперсия $s_x^2 = \frac{6}{5} * 2{,}3 = 2{,}78$, стандарттық ауытқуы $s_x = 1{,}67$.

Корреляция коэффициенті.

Тәуелділік байланысының нақты көрсеткіш өлшемі корреляция коэффициенті. Корреляция коэффициентінің екі формасы бар: теориялық және тандамалы. ρ-корреляцияның теориялық коэффициенті келесі жолмен анықталады:

$$\rho_{x,y} = \frac{\sigma_{x,y}}{\sqrt{\sigma_x^2 \sigma_y^2}}.$$

Корреляцияның теориялық коэффициенті екі кездейсоқ шаманың сызықтық байланысының тығыздығын көрсетеді және $[-1,1]$ интервалында өзгереді.

r-корреляцияның тандамалы коэффициенті теориялық дисперсияны және корреляцияны олардың ығыспаған бағаларына алмастыру арқылы анықталады:

$$r_{x,y} = \frac{\frac{n}{n-1}Cov(x,y)}{\sqrt{\frac{n}{n-1}Var(x)\frac{n}{n-1}Var(y)}} = \frac{Cov(x,y)}{\sqrt{Var(x)Var(y)}}.$$

Немесе

$$r = \frac{\sum_{i=1}^{n}(X_i - X_{opta})(Y_i - Y_{opta})}{\sqrt{\sum_{i=1}^{n}(X_i - X_{opta})^2 (Y_i - Y_{opta})^2}}.$$

Бұл жерде r – корреляция коэффиценті, n – мәліметтер саны, x_i – себепші фактор мәндері, x_{opta}, y_{opta} – орта мәні; y_i – нәтижелі фактордың мәндері, i – индекс.

Тандамалы корреляция коэффициенті х пен у сұрыптаулы мәнінің арасындағы қатаң сызықтың оң тәуелділігінде жасалатын бірлікке тең максималды мәнге ие. Аналогиялық тәсілмен r сызықтық теріс тәуелділік байланыс бар кезде, минималды мәні – 1 қабылдайды. r мәні 0 теңдігін көрсетеді r =0. Тандаулыда бақыланатын х пен у арасындағы тәуелділікте байланыс мәні жоқ.

1.4-мысал. 2000−2005 жылдарғы Қазақстанның физикалық көлемінің өсу қарқыны мен өндірістік өнімінің өсу қарқынын қарастырамыз: 15.5, 13.8, 10.5, 9.1, 10.4, 4.6. Бірінші $\bar{x}, \bar{y}$ орташаны табамыз.

$\bar{x} = 10{,}12$, $\bar{y} = 10{,}65$ дисперсиясы $s_x = 2{,}78$, $s_y = 14{,}51$ одан кейін ковариация және корреляция коэффицентін есептейміз.

$$Cov(x,y) = \frac{1}{6}(9{,}8 \cdot 15{,}5 + 13{,}5 \cdot 13{,}8 + 9{,}3 \cdot 10{,}5 +$$

$$+ 9.3 \cdot 9{,}1 + 9.4 \cdot 10.4 + 9{,}4 \cdot 4{,}6)\, 10{,}12 \cdot 10{,}65 = 2{,}5.$$

$$r_{xy} = \frac{2{,}5}{\sqrt{2{,}78 * 14{,}51}} = 0{,}39.$$

Корреляция коэффиценті анықтамасын жоспарлау арқылы, физикалық көлемнің өсу қарқыны мен өндірістік өнім көлемінің өсу қарқыны арасында байланыс тәуелділігі байқалады, бірақ өте қатты байланыс жоқ. Осы арада қатарлар санының дұрыс шешім алатындығын айта кетуге болады.

1.12-тапсырма. Кейбір бюрократикалық елде қандайда бір у индивидтің жылдық табысы *y=10000+500s+200t* формуласы-

мен анықталады, мұндағы s – тұлғаның оқыған жылдар саны; t – еңбек стажы (жыл); x – тұлға жасы (1.5-кесте) [8,10].

Төменде көрсетілген 5 жеке тұлғаны таңдау үшін $Cov(x,y)=500Cov(x,s)+200Cov(x,t)$ болатынын тексеріңіз. Аналитикалық тұрғыдан неге тең болатынын түсіндіріңіз.

1.5-кесте

Жеке тұлға	X, жасы (жыл)	S, оқыған жылдары	T, еңбек стажы	У, табысы
1	18	11	1	15700
2	29	14	6	18200
3	33	12	8	17600
4	35	16	10	20000
5	45	12	5	17000

1.13-тапсырма. 1-тапсырмадағы деректерді қолдана отырып, Var(y), Var(s) Var(t) есептеп Var(y)=250000Var(s)+ +40000Var(t)+200000Cov(s,t). Неге бұлай екенін аналитикалық тұрғыдан түсіндіріңіз.

1.14-тапсырма. Екі x және y кездейсоқ айнымалылардың бақылауы y=a +bx түзуінде орналасқан. Cov(x,y)=bVar(x) $Var(y)=b^2Var(x)$ екенін, сондай ақ, түзу қисығы оң болса, коррелияцияның таңдамалы коффициенті 1-ге тең екенін, егер түзу қисығы теріс болса, онда коффициенті 1-ге тең екенін көрсетіңіз.

Бақылау сұрақтары

1. Таңдамалы және теориялық ковариация дегеніміз не?
2. Таңдамалы дисперссия дегеніміз не?
3. Корреляция коэффициентінің қандай екі формасы бар?
4. Ковариация және дисперсия есептеу ережелері.

1.5. Жұптық регрессия

Жұптық (қарапайым) регрессия моделінде тәуелді (түсіндірілетін) айнымалының орташа мәні *x* бір тәуелсіз (түсіндіруші) айнымалының функциясы ретінде қарастырылады, яғни модельдің түрі [3,4,6,9]:

$$\hat{y}_x = f(x).$$

Сондай-ақ мұнда *y* нәтижелік мәні, ал *x* мән-фактор деп те аталады. «^» белгісі *x* және *y* айнымалылары арасында қатаң функционалдық тәуелділік жоқ екенін білдіреді. Әрбір жеке жағдайда *y* көрсеткіші екі қосылғыштан тұрады: $y = \hat{y}_x + \varepsilon$, мұндағы *y* – нәтижелік мәннің нақты мәні; $\hat{y}_x$ – регрессия теңдеуіне негізделіп табылған нәтижелік мәннің теориялық мәні; ε – регрессия теңдеуі бойынша табылған нәтижелік мәннің нақты теориялық мәнінен ауытқуын сипаттайтын кездейсоқ айнымалы.

ε кездейсоқ айнымалы қателік деп те аталады. Оған модельде ескерілмеген факторлар, кездейсоқ қателіктер, өлшеу ерекшеліктері кіреді. Оның модельде бар болуы үш көрсеткіштен туындайды: модельдің спецификациясы, кіріс мәліметтерінің таңдамалы сипаты, айнымалылардың өлшеу ерекшеліктері.

Дұрыс таңдалған модельдің спецификациясымен кездейсоқ қателіктің мөлшері байланысты: $\hat{y}_x$ нәтижелік мәннің теориялық мәні *y* нақты мәліметтеріне жақындаған сайын олар азаяды.

Спецификацияның қателеріне $\hat{y}_x$ үшін әйтеуір бір математикалық функцияның дұрыс емес таңдалуы және регрессия теңдеуінде әйтеуір бір фактордың ескерілмей қалғандығы, яғни көптік регрессия орнына жұптық регрессия қолданғандығы жатады.

Спецификацияның қателерімен бірге іріктеменің қателері де орын ала алады, олар кіріс статистикалық жиында берілгендердің біркелкі еместігіне байланысты. Егер жиын біркелкі болса, онда регрессия теңдеуінің тәжірибелік мәні болмайды. Жақсы нәтиже алу үшін жиыннан зерттелетін белгілердің аномальді мәндері бар бірліктерді алып тастайды. Және бұл жағдайда регрессияның мәндері іріктемелік сипаттамалар құрайды.

Уақыттық ақпаратты қолдану хронологиялық күндердің барлық жиынынан іріктемені құрайды. Уақыттық интервалды өзгертіп, регрессияның басқа нәтижелерін алуға болады.

Әсіресе макродеңгейде зерттегенде өлшеу қателерінің рөлі зор. Осылайша, сұраныста және тұтынуды зерттеген кезде түсіндіруші айнымалы ретінде «халық санына табыс» кең қолданылады. Осымен бірге табыс көрсеткішінің статистикалық өлшемі көптеген қиындықтармен ұштасады және оның мүмкін қателіктері болуы мүмкін, мысалы көрсетілмеген табыс болған жағдайда.

Өлшеу қателері минимумға келтірілген деп болжап, эконометрикалық зерттеулерде модельдің спецификациясының қателіктеріне ерекше көңіл аударылады.

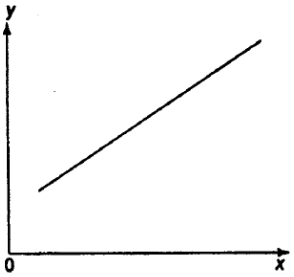
Жұптық регрессияда $\hat{y}_x = f(x)$ математикалық функцияның түрін таңдау үш әдіспен жүзеге асырылуы мүмкін [12]:

1) графикалық;

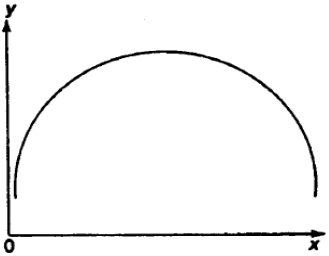
2) аналитикалық, яғни зерттелетін өзарабайланыс теориясына сүйене отырып;

3) экспериментальді.

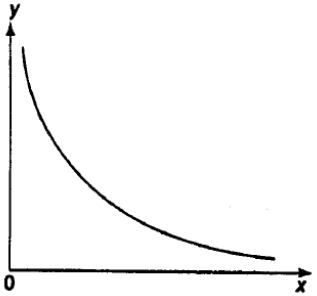
Екі белгі арасында тәуелділікті зерттеу кезінде регрессия теңдеуін таңдаудың *графикалық әдісі* көрнекті. Ол корреляция айдынына негізделген. Байланыстарды мөлшерлік бағалау кезінде қолданылатын қисықтардың негізгі түрлері 1.1-суретте келтірілген:



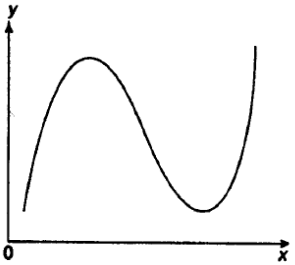
$\hat{y}_x = a + bx$ – сызықтық



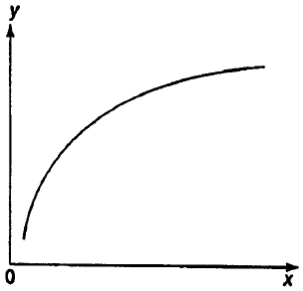
$\hat{y}_x = a + bx + cx^2$ – параболалық



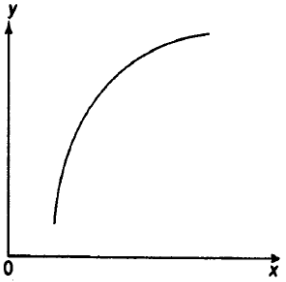
$\hat{y}_x = a + b/x$ – гиперболалық



$\hat{y}_x = a + bx + cx^2 + dx^3$ – жиынтық қисығы



$\hat{y}_x = a + x^b$ – дәрежелік



$\hat{y}_x = a + b^x$ – көрсеткіштік

1.1-сурет. Айнымалылар арасындағы байланыстарды мөлшерлік бағалау кезінде қолданылатын қисықтардың негізгі түрлері

Регрессия теңдеуінің түрін таңдауды *аналитикалық* әдіспен анықтаудың мәні ерекше. Ол зерттелетін мәндер байланысының материалды табиғатына негізделген.

Ақпаратты компьютерде өңдеу кезінде регрессия теңдеуінің түрін таңдау көбінесе *эксперименталъді әдіспен* жүзеге асырылады, яғни әртүрлі модельдерге есептелген $\sigma^2_{калд}$ қалдық дисперсияның мөлшерін салыстыру жолымен.

Егер регрессия теңдеуі корреляциялық алаңының барлық нүктелерінен өтсе, бұл тек функционалдық байланыс кезінде ғана мүмкін, бұл кезде барлық нүктелер $\hat{y}_x = f(x)$ регрессия сызығында жатқанда, $y = \hat{y}_x$ нәтижелік белгінің нақты мәндері теориялық мәндеріне сәйкес келеді, яғни олар *x* факторының әсерімен түсіндіріледі. Бұл жағдайда қалдық дисперсия $\sigma^2_{калд} = 0$.

Тәжірибелік зерттеулерде, негізінен, регрессия сызығына қатысты нүктелердің таралымы орын алады. Ол регрессия теңдеуінде ескерілмейтін факторлар әсерімен түсіндіріледі. Басқаша айтқанда, $(y - \hat{y}_x)$ нақты мәліметтердің теориялық мәліметтерден ауытқуы орын алады. Бұл ауытқулардың мөлшері осы қалдық дисперсияны есептеу негізінде жатыр:

$$\sigma^2_{\text{калд}} = \frac{1}{n}\sum (y - \hat{y}_x)^2.$$

Қалдық дисперсиясының мөлшері аз болған сайын регрессия теңдеуінде ескерілмейтін факторлар әсері аз болады.

Зерттеулер саны *x* айнымалысы кезіндегі есептелетін параметрлер санынан 7–8 есе көп болуы керек деп есептеледі. Бұл 7-ден аз зерттеу болған жағдайда сызықтық регрессияны іздеудің қажеттігі жоқ. Егер функция түрі күрделенсе, онда зерттеулер көлемін көбейту қажет болады. Яғни егер біз екінші деңгейдегі параболаны таңдасақ $\hat{y}_x = a + bx + cx^2$, онда кем дегенде 14 зерттеудің ақпарат көлемі қажет болады.

Бақылау сұрақтары

1. Жұптық сызықтық регрессия моделінің жалпы түрі қандай?
2. Жұптық сызықтық регрессия моделінде ε кездейсоқ шамасының бар болуының негізгі себептерін атаңыз?
3. Регрессиялық талдау жасау үшін қандай әдіс қолданылады?

1.6. Регрессияның жұптық сызықтық моделі және корреляция

Бұл бөлімде жұптық регрессияның қарапайым моделін – сызықтық регрессияны қарастырамыз. Сызықтық регрессия эконометрикада өзінің параметрлерінің нақты экономикалық интерпретациясы есесінен кең қолданылады [3,4,6,9].

Сызықтық регрессия

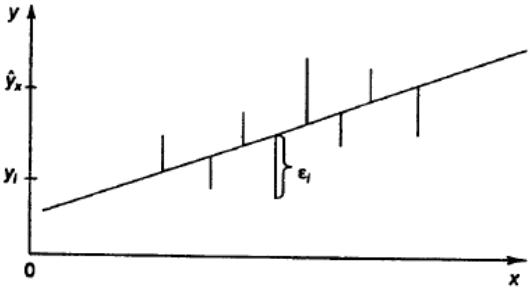
$$\hat{y}_x = a + bx \text{ немесе } y = a + bx + \varepsilon \qquad (1.8)$$

түріндегі теңдеуді табуға келеді. $\hat{y}_x = a + bx$ түріндегі теңдеу *x* факторының берілген мәндері бойынша *x* факторының нақты мәндерін қойып, нәтижелік мәннің теориялық мәндерін табуға мүмкіндік береді.

Сызықтық регрессияны құру оның параметрлерін – *a* және *b* анықтау болып табылады. Сызықтық регрессияның параметрлерін бағалаудың классикалық тәсілдемесі ең кіші квадраттар әдісіне (ЕКӘ) негізделген. ЕКӘ *a* және *b* параметрлерінің бағаларын *y* нәтижелік мәннің нақты мәндерінің $\hat{y}_x$ теориялық мәндерінен ауытқуларының квадраттарының сомасы минималды етіп алуға болады:

$$\sum_{i=1}^{n}(y_i - \hat{y}_{x_i})^2 = \sum_{i=1}^{n}\varepsilon_i^2 \to \min.$$

Яғни графикте сызықтардың барлық жиынынан регрессия сызығы вертикаль бойынша бұл сызық пен нүктелер арасындағы арақашықтықтар квадраттар сомасы минималды болатындай болып таңдалады (1.2-сурет) [7]:



1.2-сурет. Қалдықтардың минималды дисперсиясы бар регрессия сызығы

Математикалық талдау курсынан білетініміз: (1.9) функцияның минимумын табу үшін *a* және *b* параметрлерінің әрқайсысы бойынша жеке туынды тауып, оларды нөлге теңестіреміз. $\sum_i \varepsilon_i^2$ -ні *(Sa,b)* арқылы белгілейміз, онда:

$$S(a,b) = \sum (y - a - bx)^2;$$

$$\begin{cases} \dfrac{dS}{da} = -2\sum (y - a - bx) = 0; \\ \dfrac{dS}{db} = -2\sum x(y - a - bx) = 0. \end{cases}$$

Қиын емес түрлендірулерден кейін, *a* және *b* параметрлерін бағалау үшін келесі сызықтық теңдеулер жүйесін аламыз:

$$\begin{cases} an + b\sum x = \sum y \\ a\sum x + b\sum x^2 = \sum xy \end{cases}$$

(1.11) теңдеулер жүйесін шешіп, *a* және *b* параметрлерінің жанама бағаларын табамыз. Келесі дайын формулаларды қолданамыз, олар (1.11) теңдеулер жүйесін шешуден шығады:

$$a = \bar{y} - b\bar{x},\ \ b = \frac{\mathrm{cov}(x, y)}{\sigma_x^2},$$

мұндағы $\mathrm{cov}(x, y) = \overline{xy} - \bar{x} \cdot \bar{y}$ – x және y мәндердің ковариациясы, $\sigma_x^2 = \overline{x^2} - \bar{x}^2$ – x мәнінің дисперсиясы және $\bar{x} = \frac{1}{n}\sum x$, $\bar{y} = \frac{1}{n}\sum y$, $\overline{yx} = \frac{1}{n}\sum yx$, $\overline{x^2} = \frac{1}{n}\sum x^2$.

b параметрі *регрессия коэффициенті* деп аталады. Оның мөлшері фактордың бір бірлікке өзгерген кездегі нәтиженің орташа өзгерісін көрсетеді.

Регрессия коэффициентінің нақты экономикалық интерпретациясының мүмкіндігі регрессияның сызықтық теңдеуін эконометрикалық зерттеулерде кең тараған етті.

Қағаз жүзінде *a – y-тің* $x = 0$ кезіндегі мәні. Егер *x* мән-фактордың нөлдік мәні жоқ болса, онда жоғарыда көрсетілген трактовканың мәні жоғалады, яғни *a* параметрінің экономикалық мәні жоқ.

Регрессия теңдеуі байланыс тығыздығының көрсеткішімен толықтырылады. Мұндай көрсеткіш ретінде сызықтық регрессияны қолданған кезде r_{xy} корреляция сызықтық коэффициенті қолданылады.

Сызықтық функцияны таңдау сапасын бағалау үшін *детерминация коэффициенті* деп аталатын r_{xy}^2 сызықтық корреляция коэффициентінің квадраты есептеледі. Детерминация коэффициенті нәтижелік мәннің жалпы дисперсиясында регрессиямен түсіндірілетін *y* нәтижелік мәннің дисперсияның үлесін сипаттайды:

$$r_{xy}^2 = \frac{\sigma_{\text{факт}}^2}{\sigma_{\text{y}}^2} = 1 - \frac{\sigma_{калд}^2}{\sigma_y^2},$$

мұнда $\sigma_y^2 = \frac{1}{n}\sum (y - \bar{y})^2$, $\sigma_{факт}^2 = \frac{1}{n}\sum (y_x - \bar{y})^2$, $\sigma_{калд}^2 = \frac{1}{n}\sum (y - \hat{y}_x)^2$.

$1-r_{xy}^2$ көрсеткіші модельде ескерілмей қалған факторлар әсерінен пайда болған *y* дисперсияның үлесін көрсетеді.

Сызықтық регрессияның теңдеуі табылғаннан кейін теңдеудің және оның параметрлердің мәнділін бағалануы жасалады.

Әрбір зерттеу бойынша қатысты ауытқулардан тұратын модельдің сапасы туралы жалпы пікір алу үшін, аппроксимацияның орташа қатесі анықталады:

$$\overline{A} = \frac{1}{n}\sum \left| \frac{y - \hat{y}_x}{y} \right| * 100\%.$$

Аппроксимацияның орташа қатесі 8–10 % аспауы керек. Регрессия теңдеуінің мәнділігін бағалау Фишерді *F*-критерийі негізінде жүзеге асырылады. Дисперсиялық талдаудың негізгі ойына сәйкес, *y* айнымалысының $\overline{y}$ орташа мәнінен ауытқу квадратының сомасы екі бөлікке бөлінеді – «түсіндірілген» және «түсіндірілмеген»:

$$\sum (y - \overline{y})^2 = \sum (\hat{y}_x - \overline{y})^2 + \sum (y - \hat{y}_x)^2,$$

мұндағы $\sum (y - \overline{y})^2$ – ауытқу квадраттарының сомасы; $\sum (\hat{y}_x - \overline{y})^2$ – регрессиямен түсіндірілген ауытқу сомасы; $\sum (y - \hat{y}_x)^2$ – модельде ескерілмей қалған факторлар әсерін сипаттайтын ауытқу квадраттарының қалдық сомасы.

Дисперсиялық талдау схемасы 1.6-кестеде келтірілген.

(*n* – зерттеулер саны, *m* – *x* айнымалысы кезіндегі параметрлер саны).

1.6-кесте

Дисперсиялық талдау схемасы

Дисперсия компоненттері	Квадраттар сомасы	Еркіндік дәрежелерінің саны	Бір еркіндік дәрежесіне шаққандағы дисперсия
Жалпы	$\sum(y-\bar{y})^2$	n-1	$S^2_{жалп}=\frac{\sum(y-\bar{y})^2}{n-1}$
Факторлық	$\sum(\hat{y}_x-\bar{y})^2$	m	$S^2_{жалп}=\frac{\sum(\hat{y}_x-\bar{y})^2}{m}$
Калдық	$\sum(y-\hat{y}_x)^2$	n-m-1	$S^2_{калд}=\frac{\sum(y-\hat{y}_x)^2}{n-m-1}$

Дисперсиядағы бір дәрежелі еркіндікті анықтау салыстырмалы түрдегі дисперсияға алып келеді. Факторлық және қалдық дисперсияны бір дәрежелі еркіндікте есептеп, салыстыра отырып, Фишердің *F*-критерийінің көрсеткішін аламыз:

$$F=\frac{S^2_{\text{факторлык}}}{S^2_{\text{калд}}}$$

Фишердің *F*-критерийінің нақты мәні (1.14) $k_1=m$ және еркіндік дәрежесі $k_2=n-m-1$ кезінде α мәнділік деңгейінде кестелік $F_{кесте}(\alpha;\ k_1;k_2)$ мәнімен салыстырылады. Бұл кезде *F*-критерийдің нақты мәні кестелік мәннен артық болса, онда теңдеудің статистикалық мәні расталады [2,3,11].

Жұптық сызықтық регрессия үшін $m=1$, сондықтан

$$F=\frac{S^2_{\text{факторлык}}}{S^2_{\text{калд}}}=\frac{\sum(\hat{y}_x-\bar{y})^2}{\sum(y-\hat{y}_x)^2}(n-2).$$

Қосымша: Фишер $F_{нақты}$ және Стьюдент мәндерін тексеру кестесі [2,3,11].

F-критерийдің көрсеткіші r_{xy}^2 детерминация коэффициентімен байланысты, оны келесі формуламен анықтауға болады:

$$F = \frac{r_{xy}^2}{1 - r_{xy}^2}(n-2).$$

Жұптық сызықтық регрессияда теңдеудің және оның параметрлерінің мәнділігі бағаланады. Осы мақсатпен параметрлердің әрқайсысы бойынша оның стандарттық қатесі анықталады: m_b және m_a.

Регрессия коэффициентінің стандарттық қатесі келесі формуламен анықталады:

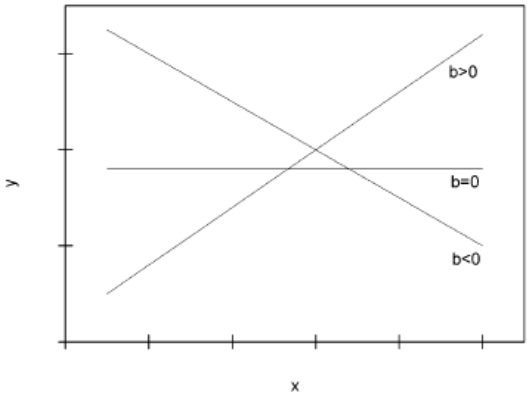
$$m_b = \sqrt{\frac{S_{\text{калд}}^2}{\sum(\text{x} - \bar{\text{x}})^2}} = \frac{S_{\text{калд}}}{\sigma_{\text{x}}\sqrt{n}},$$

мұндағы $S_{калд}^2 = \frac{\sum(y - \hat{y}_x)^2}{n-2}$ – еркіндіктің бір дәрежесіне қалдық дисперсиясы.

Стандарттық қате көрсеткіші және еркіндік дәрежесі *n*-2 кезінде Стьюденттің *t*-тарату регрессияның коэффициенттерінің орнықтылығын тексеру үшін және оның сенімділік интервалын есептеу үшін қолданылады.

Регрессияның коэффициенттерінің орнықтылығын бағалау үшін оның көрсеткіші оның стандарттық қатесімен салыстырылады, яғни Стьюденттің *t* – критерийінің нақты мәні анықталады: $t_b = \frac{b}{m_b}$, ол сосын еркіндік дәрежесі *n*-2 және мәнділіктің белгілі деңгейінде кестелік мәнімен салыстырылады. Регрессия коэффициенті үшін сенімділік интервалы $b \pm t_{кесте} m_b$ ретінде анықталады. Регрессия коэффициентінің таңбасы *x* ($b>0$) мән-фактордың өсуі кезінде *y* нәтижелік мәннің өсуін, ($b<0$) мән-фактордың өсуі кезінде нәтижелік мәннің азаюын немесе оның тәуелсіз айнымалыдан тәуелсіздігін ($b=0$) көрсетеді (1.3-сурет), онда Регрессия

коэффициенті үшін сенімділік интервалының шектерінің қарсы нәтижелері болмауы керек, мысалы, $1,5 \le b \le 0,8$. Осындай түрдегі жазба регрессия коэффициентінің ақиқат мәнінің оң және теріс көрсеткіштері, және нөл бар екенін көрсетеді, ал бұл мүмкін емес [12].



1.3-сурет. *b* параметрінің мәніне байланысты регрессия сызығының орналасуы.

a параметрінің стандарттық қатесі келесі формуламен анықталады:

$$m_a = \sqrt{S_{\text{калд}}^2 \frac{\sum x^2}{n\sum (x-\bar{x})^2}} = S_{\text{калд}} \frac{\sqrt{\sum x^2}}{\sigma_x n}.$$

Берілген параметрдің шындылығын бағалау процедурасы регрессия коэффициенті үшін қарастырылған процедурадан айырмашылығы жоқ.

t-критерий есептеледі: $t_a = \frac{a}{m_a}$, оның көрсеткіші еркіндік дәрежесі n-2 кезіндегі кестелік мәнімен салыстырылады.

Корреляцияның сызықтық коэффициентінің мәнділігі m_r корреляция коэффициентінің көрсеткіші негізінде тексеріледі:

$$m_r = \sqrt{\frac{1-r^2}{n-2}}.$$

Стьюденттің *t*-критерийінің нақты мәні келесідей анықталады: $t_r = \frac{r}{m_r}$.

Стьюденттің *t*-критерийі және Фишердің *F*-критерийі арасында байланыс бар:

$$|t_b| = |t_r| = \sqrt{F}.$$

Болжау есептерінде регрессияның теңдеуі бойынша $x = x_0$ кезінде нүктелік болжау ретінде y_0 жеке мәні анықталады, яғни $\hat{y}_x = a + bx$ сызықтық теңдеуіне тиісті *x* мәнін қою жолымен. Бірақ нүктелік болжау мүмкін емес, сондықтан ол стандарттық қатені есептеумен

$$m_{\hat{y}0} = \sqrt{S^2_{\text{калд}}\left(1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{\sum (x - \bar{x})^2}\right)} = \sqrt{S^2_{\text{калд}}\left(1 + \frac{1}{n} + \frac{(x_\rho - \bar{x})^2}{n\sigma_x^2}\right)},$$

мұнда $S^2_{калд} = \frac{\sum (y - \hat{y}_x)}{n-2}$ және y_0^* болжау мәнінің *сенімділік интервалын* құрумен толықтырылады:

$$\hat{y}_0 - m_{\hat{y}0} t_{\text{кесте}} \le y_0^* \le \hat{y}_x + m_{\hat{y}_x} t_{\text{кесте}}.$$

Бақылау сұрақтары

1. Регрессиялық талдау мәні неде?
2. Регрессияның бағаланған таңдеуінің a және b көрсеткіштері қандай формуламен есептелінеді?
3. Детерминация коэффициенті нені сипаттайды?
4. Детерминация коэффициенті қандай мәндерді қабылдай алады?

5. Теңдеудің және оның арнайы параметрлерінің шындылығын бағалау (Фишердің *F*-критерийі және Стьюденттің *t*-критерийі).

6. Регрессияның сызықтық теңдеу бойынша болжау. Аппроксимацияның орташа қатесі.

7. Регрессиялық талдау есебінің мәні неде?

8. Сызықтық функцияның сапасын бағалау үшін қандай коэффициент қолданылады?

1.7. Қалыпты үлестірімді шаманың параметрлерін интервалдық бағалау

$N(a,\sigma)$ қалыпты заңы бойынша $x_1, x_2,...,x_n$ таңдамасын қарастырамыз. a мен σ параметрлері үшін бағалар сәйкес $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ және

$$S_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \qquad (1.22)$$

болсын. $\frac{1}{\sigma}(\bar{x} - a)\sqrt{n}$ кездейсоқ шамасының үлестірімі N(0,1) қалыпты заң болады. $\frac{1}{\sigma}s^2$ кездейсоқ шамасы $\frac{1}{n-1}\chi^2$ бойынша үлестірімді. Демек, $t_{n-1} = \frac{\frac{1}{\sigma}(\bar{x} - a)\sqrt{n}}{\sqrt{\frac{1}{\sigma^2}s^2}} = \frac{(\bar{x} - a)\sqrt{n}}{s}$ кездейсоқ шамасы Стьюдент заңы бойынша үлестірімді, мұндағы еріктілік дәрежесі $(n-1)$-ге тең. Стьюдент үлестірімінің параметрлерге тәуелсіз екендігі белгілі. Міне, бұл – қалыпты заңның негізгі қасиеттерінің бірі [1,3,9].

Математикалық a үміті үшін аралық баға құруға көшелік. Стьюдент үлестірімнің кестесі бойынша берілген $\boldsymbol{\alpha}$ саны арқылы $P\left|t_{n-1}\right| \le t_\alpha = 1-\alpha$ болатындай етіп, t_α санын табамыз. Содан

кейін t_{n-1}-дің орнына $\left(\frac{1}{s}(\bar{x}-a)\sqrt{n}\right)$ өрнегін қойып, пайда болған теңсіздікті a-ға қатысты шешеміз. Сонда a үшін

$$P\left(\bar{x}-\frac{s}{\sqrt{n}}t_{\alpha}\le a\le \bar{x}+\frac{s}{\sqrt{n}}t_{\alpha}\right)=1-\alpha \qquad (1.23)$$

аралық бағалауы келіп шығады. Егер, мәселен, бұл формулаға қатысты t_a санын, $\sigma = s$ деп алып, қалыпты заңның кестесі бойынша есептесек, онда t_a саны Стьюдент заңы бойынша есептеген мәніне қарағанда аз болар еді. Демек, a үшін (1.22) аралық баға (1.23) аралық бағамен салыстырғанда «ауқымды» болады.

Енді σ^2 дисперсиясын бағалауға көшелік. $\frac{n-1}{\sigma^2}s^2$ кездейсоқ шамасы χ^2_{n-1} (хи – квадрат) үлестірімі. Берілген α бойынша $q_1(\alpha)$ және $q_2(\alpha)$ сандарын $P\left(q_1(t)\le \chi^2_{n-1}\prec q_2(t)=\alpha\right)$ болатындай етіп, тауып аламыз. Бұл шартты қанағаттандыратын q_1 ***(α)*** мен $q_2(\alpha)$ бірмәнді болып табылмайды. Сондықтан бұл сандарды, әдетте,

$$P\left(\chi^2_{n-1}\le q_1(\alpha)\right)=\frac{\alpha}{2}, \qquad P\left(\chi^2_{n-1}\ge q_2(\alpha)\right)=\frac{\alpha}{2},$$

болатындай етіп, сайлап алады. Содан кейін χ^2_{n-1}-тың орнына $\frac{(n-1)s^2}{\sigma^2}$ өрнегін қойып, σ^2 арқылы теңсіздікті шешсек, σ^2 дисперсиясы үшін мынандай аралық баға келіп шығады:

$$P\left(\frac{(n-1)s^2}{q_2(\alpha)}\le\sigma^2\le\frac{(n-1)s^2}{q_1(\alpha)}\right)=1-\alpha. \qquad (1.24)$$

«Қателер теориясы» деп аталып жүрген есептеулердің негізгі мазмұны (1.23) және (1.24) формулалар арқылы анықталады.

$N(\alpha,\sigma)$ қалыпты заңынан алынған $x_1,x_2,...,x_n$ таңдамасы берілсін. a_0 саны алдын ала берілген болсын. «*a* математикалық үміті a_0 санына тең» деген болжам дұрыс па? Міне, осы сұраққа жауап берелік. $a = a_0$ болжамын H_0 арқылы белгілелік. H_0 болжамының бірнеше альтернативасы (мүмкін болатын екі ұйғарымның бірін ғана таңдап алуға мәжбүрлік) болуы мүмкін. Мынадай екі альтернатива жиі кездеседі:

а) $a>a_0$, яғни егер $a = a_0$ дұрыс болмаса, онда $a>a_0$ (теңсіздік таңбасын өзгерткенде бұл альтернатива $a<a_0$ альтернативасына көшеді);

б) $a\neq a_0$, яғни егер $a = a_0$ дұрыс болмаса, $a<a_0$ және $a>a_0$ жағдайларының екеуі де болуы мүмкін.

а) альтернативасы *бір жақты деп, ал;*

б) альтернативасы *екі жақты деп* аталады.

a математикалық үміті үшін баға етіп $\bar{x}$ -ті аламыз. Берілген α мәнділік деңгейі бойынша, қайсыбір C_1 мен C_2 сандары сайланып алынған делік. Сонда: егер $\bar{x} - a_0 > C_1$ болса, онда а) альтернативасында H_0 болжамын қабыл алмау керек, ал $\left|\bar{x} - a_0\right| > C_2$ болса, онда а) альтернативасында H_0 болжамын қабыл алмау керек.

Енді C_1 мен C_2 сандарын сайлап алу мәселесін қарастырайық. Егер H_0 болжамы шындық болса, онда $\alpha_{n-1} = \frac{1}{s}(\bar{x} - a_0)\sqrt{n}$ кездейсоқ шаманың үлестірімі Стьюдент заңы болады. Кесте бойынша $P(t_{n-1} \le d_\alpha) = \alpha$ болатындай етіп, d_a санын табамыз. Сонда шындық H_0 болжамында

$$P\left(\bar{x} - a \succ \frac{sd}{\sqrt{n}}\right) = \alpha$$

басқаша айтқанда, $C_1 = \frac{sd_2^1}{\sqrt{n}}$. Сонымен C_1 санын сайлап алу, тәжірибе саны *n*-ге және әрбір тәжірибе дәлдігі үшін s сипаттамасына байланысты.

б) альтернативасы жағдайында *d*-ның орнына $P\left(\left|t_{n-1}\right| \succ d_2 = \alpha\right)$ қатынасы қанағаттанатындай d_2 санын алу керек. Ал Стьюдент үлестірімінің анықтамасынан кез келген *z* үшін $P(|t_{n-1}|\geq z) = 2P(t_{n-1} \geq z)$ теңдігі келіп шығады. Сондықтан да $d_{\alpha/2}=d'_{\alpha}d_{\alpha}$ шамасын *біржақты сенімді шекаралық* деп, ал d'_{α} шамасын *екіжақты сенімді шекаралық* деп атайды.

1.15-тапсырма. Интегралды есептеу әдісі көмегімен өнімнің бірден түсуі және қорды бірқалыпты жұмсау кезінде қордың орташа мөлшерін анықтаныз.

Бақылау сұрақтары

1. Қандай интервал сенімді деп аталады?
2. Сіз қандай екі гипотезаны білесіз?
3. Гипотезаларды қандай мақсаттар үшін тұрғызады?

1.8. Жұптық сызықтық регрессия параметрлерін есептеу

1.5-мысал. Ферма мамық жүнді аңдарды асыраумен айналысады. Маңызды талдама негізінде фермерде барлық технологиялық нормативтер және тамақтану нормасы сақталып тұр. Сонда аң салмағы оның жасына тәуелді болуы керек екендігін регрессиялық анализ арқылы негіздеу керек [5,6,11].

Мамық жүнді аңның салмағының оның жасына сандық тәуелділігін анықтайық. У – оның ай сайынғы салмағы (кг), Х – оның айлық жасы (ай) (1.7-кесте).

1.7-кесте

Бастапқы мәліметтер

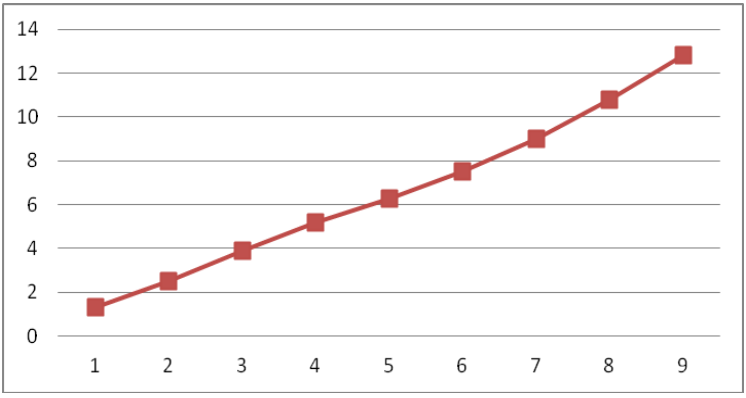
Xi – жасы, ай	0	1	2	3	4	5	6	7	8
Yi – салмағы, кг	1.3	2.5	3.9	5.2	6.3	7.5	9	10.8	12,8

Тапсырма:

1. Факторлар байланысының графигін сыз;
2. Тәуелділік тығыздығын анықта;
3. У(х) – жұптық регрессияның теңдеуін құрастыр;
4. Регрессия теңдеуінің параметрлерін анықта;
5. Регрессия теңдеуінің адекваттылығын тексер;
6. Регрессия параметрлерінің статистикалық маңыздылығын бағала;
7. Регрессия параметрлерінің сенімділік интервалын анықта;
8. У болжауын тап;
9. Болжаудың дәлділігін бағала.

Шешімі. Есептеулерді кестелік түрде Excel-де келесі 1.8-кестеде көрсетейік.

1. Есептің берілген мәндері бойынша координаттар жүйесінде графигін сызамыз (1.4-сурет). Мәліметтердің графигі түзу сызықтың бойында орналасқандығы корреляция өрісінің жинақтылығын көрсетеді, сонымен қатар корреляция коэффициентінің бірге жақын болатындығына кепіл.



1.4-сурет. Жұптық регрессия графигі

1.8-кесте

Жұптық регрессияның есептеулері

№	Xi	Yi	Xi-Xорт	Yi-Yорт	(Xi-Xорт) (Yi-Yорт)	(Xi-Xорт)^2	(Yi-Yорт)^2	Xi^2	Xi*Yi	Y*	Ai	(Y-Y*)2
1	2	3	4	5	6	7	8	9	10	11	12	13
1	0	1,3	-4	-5,29	21,16	16	27,97	0	0	1,03	0,21	0,074
2	1	2,5	-3	-4,09	12,27	9	16,72	1	2,5	2,42	0,03	0,007
3	2	3,9	-2	-2,69	5,38	4	7,23	4	7,8	3,81	0,02	0,008
4	3	5,2	-1	-1,39	1,39	1	1,93	9	15,6	5,20	0,00	0,00
5	4	6,3	0	-0,29	0,00	0	0,08	16	25,2	6,59	0,05	0,083
6	5	7,5	1	0,91	0,91	1	0,83	25	37,5	7,98	0,06	0,229
7	6	9	2	2,41	4,82	4	5,81	36	54	9,37	0,04	0,136
8	7	11	3	4,21	12,63	9	17,73	49	75,6	10,76	0,00	0,002
9	8	13	4	6,21	24,84	16	38,58	64	102,4	12,15	0,05	0,424
Бар-лығы	36	59	0	0,00	83,40	60	116,9	204	320,60	59,3	0,47	0,96
Орта-ша	4	6,59	0	0,00	9,27	6,67	12,99	22,67	35,62	6,59		0,11
Пара-метр		a=	**1,03**	b=	**1,39**	r=	**0,9959**	R^2=	**0,9918**	Aорт=	**5,22%**	

R^2= 0,9917624. F=842,7577. u =29,03029

2. Байланыс тығыздығын орнату үшін *r* корреляция коэффициентінің мәнін табамыз. Ол үшін 6–8 графалардың сомалы мәндерін пайдаланамыз. Сонда

$$r = \frac{83{,}4}{\sqrt{60 * 116{,}8899}} = 0{,}9959.$$

Байланыс өте тығыз және тура. Корреляция коэффициенті 1-ге өте жақын. Детерминация коэффициентін анықтайық $(r)^2=(0.9959)^2=0.9818$. Нәтижелі фактордың вариациясы 98,18 %-ға х фактордың вариациясымен түсіндіріледі, ал 2 % өзгеруі ескерілмеген факторларға келеді. Егер таңдалған факторлар арасында тығыз байланыс болса, онда регрессия теңдеуін құруға болады.

3. Функция түрін анықтау үшін у-тің х-ке тәуелді болып келетін тәуелділік графигіне қараймыз (1.4-сурет). Суреттен

нүктелер түзу сызықты бойлай орналасқаны көрініп тұр. Яғни, сызықтық функцияны таңдаймыз.

Регрессия теңдеуі мынадай түрден тұрады y=a+b*x.

4. a және b параметрлерін анықтау үшін мына формуланы қолданамыз

$$\begin{cases} a*n + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \\ a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i\, y \end{cases}$$

Кестенің 2–3, 9–10 графалардағы шыққан сомалы мәндерді қолдана отырып, тендеу жүйесін аламыз

$$9a+36b=59,3;$$
$$36a+204b=320,6.$$

Бұдан *a=1,0289, b=1,39*, сонда регрессия тендеуі y=1,0289+1,39x. Аңның жасы 1 айға өскен сайын оның орташа салмағы 1,39 кг көбейіп жатыр.

5. Регрессия теңдеуінің адекваттылығы $A_{орт}$, $t_{факт}$, F мәндерінің шартты қанағаттыруына байланысты.

А) Аппроксимацияның орташа қателігін табайық, ол үшін 12-графаны қолданамыз.

$$A_{орта} = \sum \frac{A_i}{n} = \frac{0,4699}{9} * 100\% = 5,2215.$$

Алынған $A_{орта}$ мәні 6 % аспағандықтан, регрессия теңдеуінің сызықты функция арқылы өрнектелуі адекватты деп айтуға болады.

Б) Регрессия моделінің адекваттылығы Фишер $F_{нақты}$ және Стьюдент критерийі Х көмегімен 13 жүргіземіз[2,3,11].

$$F = \frac{r^2}{1-r^2}(n-2) = 842,75.$$

$F_{кесте}$ = 5,12<$F_{факт}$ =842,75 кездейсоқ фактордың гипотезасы ауытқиды

Стьюдент критерийін келесі формула бойынша шығарамыз $t_{кесте} = \sqrt{F} = 29,03$. Яғни

$$t_{кесте}=2,26<t_{факт}=29,03.$$

Корреляция коэффициентінің Стьюдент критерийінің нақты мәні $t_{кесте}=2,26 < t_{факт}=846,25$ арақатынасы y және x арасында байланыстың кездейсоқ емес екенін білдіреді. Басқа барлық факторлар тұрақты болып қалатын болса, яғни, барлық технологиялық нормативтер орындалса, онда аңның салмағы жасына орнықты тәуелді.

Мұнда регрессия теңдеуі мәліметтерге сәйкес, яғни, алынған теңдеу *y* және x факторларын сандық тәуелділігін анық сипаттайды.

6. Регрессия моделі параметрлерінің статистикалық маңыздылығын бағалауды келесі формула бойынша жүргіземіз:

$$t_b = \frac{b}{m_b}; \qquad t_a = \frac{b}{m_a}.$$

Сызықтық регрессияның кездейсоқ қателіктері келесі формула бойынша анықталады

$$m_a = \sqrt{\frac{\sum (Y - Y^*)^2 \sum X^2}{n(n-2)\sqrt{(X - X_{орт})^2}}}; \qquad t_a = \frac{b}{m_a};$$

$$m_b = \sqrt{\frac{\sum (Y - Y^*)^2 \sum X^2}{(n-2)\sqrt{(X - X_{орт})^2}}}; \qquad t_b = \frac{b}{m_b}.$$

m_a есептеу үшін 1.8-кестедегі 7,9,13 графаларды қолданамыз.

$m_a=0,2279$, $m_b=0.04788$ бұдан

$t_a=4,5134>2,26>t_{кесте}$;

$t_b=29,03>2.26\ t_{кесте}$.

Алынған регрессия тендеуінің параметрлерін статистикалық бағалаулары аңның салмағы оның жасына статитикалық түрде тәуелді екенін бекітуге мүмкіндік береді және олар орнықты тәуелді екенін көрсетеді.

7. Регрессия параметрлерінің сенімділік интервалы келесі формула бойынша есептеледі $a + t_{кесте}*m_a$ және $b \pm t_{кесте}*m_b$. Ол үшін әр параметрлердің Δ шекті қателігін анықтаймыз

$$\Delta_a = t_{кесте} m_a = 2,26 * 0,2279 = 0,515;$$

$$\Delta_b = t_{кесте} m_b = 2{,}26 * 0{,}04788 = 0{,}1082.$$

Сенімді интервалды есептеу формуласы келесі түрде берілген:

$$\gamma_a = a \pm \Delta_a = 1{,}0289 \pm 0{,}515;$$

$$\gamma_{a\min} = a - \Delta_a = 1{,}0289 - 0{,}515 = 0{,}5139;$$

$$\gamma_{a\max} = a + \Delta_a = 1{,}028 + 0{,}515 = 1{,}5439.$$

Онда *a* параметрі интервалда болады 0,5139<a<1,5439.

b параметрін солай есептейміз

$$\gamma_b = b \pm \Delta_b = 1{,}39 \pm 0{,}1082;$$

$$\gamma_b = b - \Delta_b = 1{,}39 - 0{,}1082 = 1{,}2818;$$

$$\gamma_b = b + \Delta_b = 1{,}39 + 0{,}1082 = 1{,}4982.$$

Сонда b параметрі мына интервалда болады 1,2818<b<1,4982.

Егер сенімді интервал шекарасына нөл түссе, яғни, төменгі шекара теріс, ал жоғарғы шекара оң болса, онда бағаланып отырған параметр нөлдік деп қабылданады. Өйткені, ол бір уақытта екі мәнді, теріс және оң мәндерді қабылдай алмайды.

Осыдан кейін алынған регрессия теңдеуін болжау үшін қолдануға болады.

8. $У_{болж}$ болжаудың мәнді регрессия тендеуіндегі У=0,87+1,41*х себепші х фактордың жоспарлы немесе болжау мәнін қою арқылы анықтаймыз. Осы есепті х фактор айдың ретін білдіргеніне қарай келесі ай тоғызыншы болады, соңда $х_{жосп}$=9.

Мәліметтерді жинау уақыт периоды бойынша жинақталған, яғни, х болжау мәні келесі период болады. Мысалы, $Х_{болж}$=9.

$У_{болж}$=1,0289+1,39*9= 13,54. Бұл дегеніміз 1 айдан соң аңның орташа салмағы 13,54 кг болатынын білдіреді. Ал 2 айдан соң салмағы

$У_{болж}$=1,0289+1,39*10= 14,93 кг болады.

9. Болжаудың орындалу аралығын табу үшін Стандарттық қателіктерді есептейміз.

1.6-мысал. 10 периодтағы болжаудың орташа тұрақты қателігі келесі формула бойынша есептеледі: мұнда $m(Y_{болж})$ = 0,4053 және болжаудың сенімді интервалын құрастырамыз

$$\gamma_y = У_{болж} \pm \Delta_{болж} = Y_{болж} \pm \Delta_y ,$$

$$\gamma_{y_{min}} = У_{болж} - t_{кесте} * m\,(Y_{болж}) = 14{,}93 - 2.26 * 0{,}4053 = 14{,}01,$$

$$\gamma_{y_{max}} = Y_{болж} + t_{кесте} * m(Y_{болж}) = 14{,}93 + 2{,}26 * 0{,}484 = 15{,}84.$$

Осылайша, $У_{болж}$ [14,01;15,84] интервалында болады. Сонда келесі екі айдан кейін, яғни он айдан кейін аңның салмағы орташа 14,93 кг болады, немесе 14,01кг және 15,84кг аралықта болады.

1.16-тапсырма. Берілгені:

Тауар айналымы, млн. тг.	1	4	8	12	15
Жарнамаға кеткен шығын, %	4	8,2	2	3,2	4,3

Тауар айналымы мен жарнамаға кеткен шығын арасындағы тәуелділікті анықтау. График тұрғызу. Алынған нәтижені түсіндіру.

1.17-тапсырма. Тауар қорының мөлшері мен сатылым көлемі арасындағы тәуелділікті анықтау, егер:

Тауар қоры, мың.тг	430	375	235	330	270	301	375
Сатылым көлемі, мың тг/де, тәулік	40,0	33	19	22	23,5	21	32

Логистика тұрғысынан нәтижені түсіндіру.

1.18-тапсырма. Қадағалау 1.9-кестесі бойынша

1.9-кесте

i	1	2	3
X	4	5	9
Y	-5	8	11

q=-10.667+2.5*z қатар желілік регрессия бойынша теңеу құрылған. R^2 детерминация коэффициентінің мәнін тауып, оның мәніне F – статистиканың көмегімен баға беріңдер.

1.19-тапсырма. Алдағы берілген есеп бойынша өндіріс көлемінің күтпелі мөлшеріне болжам жасаңыз. Егер фирма 10 жұмысшыны жалдаса, онда оның өндіріс көлемінің күтімін және жеке мәніне 95 %-дық сенімді интервал құрыңыз.

1.9. Өзара тәуелділік байланысын зерттеудегі корреляциялық-регрессиялық әдісімен модельдеу

Компьютердегі ҚБП Excel программасы арқылы регрессиялық талдау оңай және жылдам орындалады. Регрессиялық талдауды жүргізу үшін арнаулы программалар пакетін қолданамыз (*1.5-мысал*). Бұл үшін келесі қадамдарды орындайық:

1. Алғашқы мәліметтерді енгіземіз (1.10-кесте)

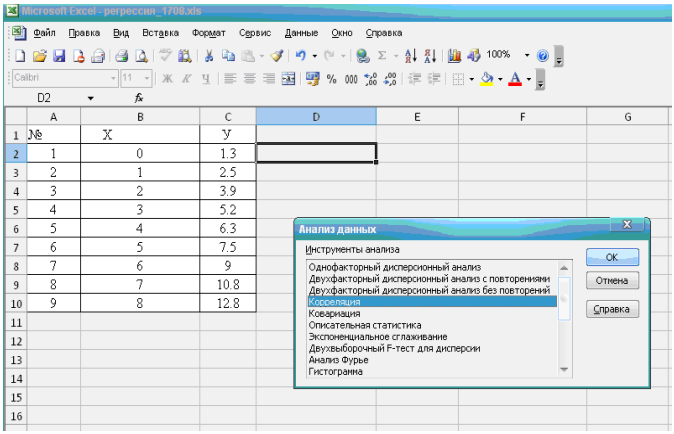
1.10-кесте

Бастапқы мәліметтер

№	X	У
1	0	1,3
2	1	2,5
3	2	3,9
4	3	5,2
5	4	6,3
6	5	7,5
7	6	9
8	7	10,8
9	8	12,8

2. Excel ортасында бастапқы мәліметтер кестесін құрайық;

3. Негізгі мәзірден мына пункттерді таңдаңыз: Сервис/Анализ данных/Корреляция/Ок

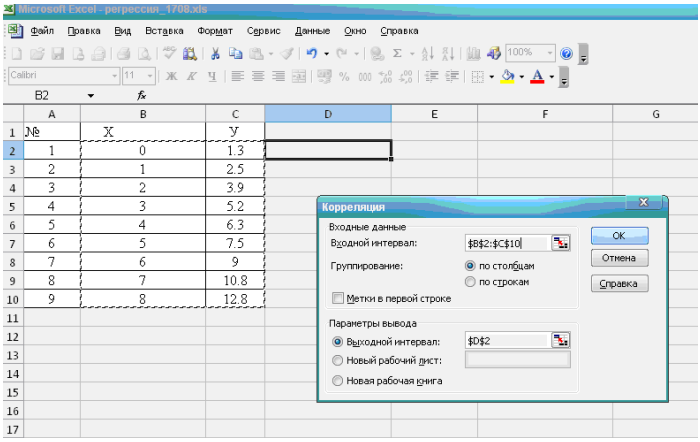


1.5-сурет . «Корреляция» капшығын шығару

3. Диалогтық терезенің және шығару параметрлерін толтырыңыз;

Кіргізу интервалы; X және У мәндерінен тұратын барлық бағаналарды белгілеңіз; B2;C10; OK.

Шығару интервалы: нәтижені шығару үшін бос ұяшық аймағын белгілеңіз, мысалы: D2;OK.(1.6-сурет)



1.6-сурет. Диалогтық терезе

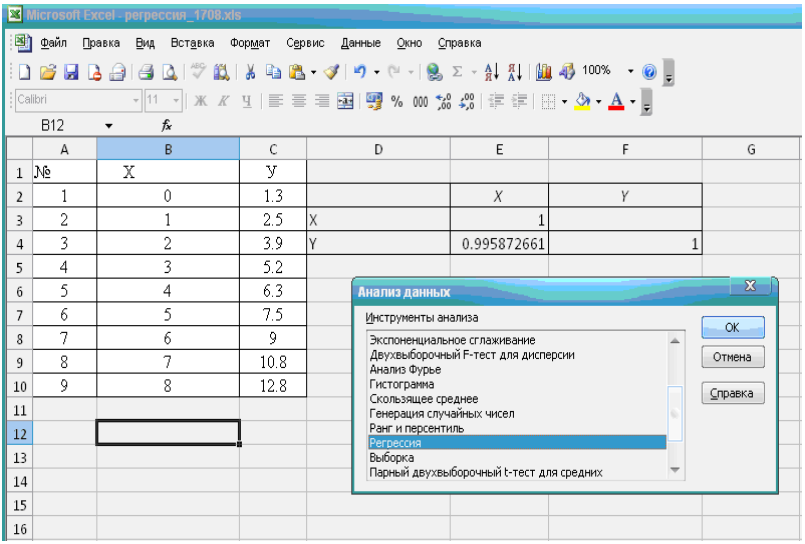
Excel У және Х жұптық корреляцияларының коэффициенттері кестесін көрсетеді: (1.11-кесте)

1.11-кесте. Компьютерде шығарылған есептің шешу нәтижелері

	X	*Y*
X	1	
Y	0,995872661	1

r_{yx}=0,9959 у және х аралығындағы корреляция коэффициентін көрсетеді.

Регрессия теңдеуінің параметрлерін есептеу үшін мәліметтерді талдау **Регрессия** пунктін қолданамыз [9,11]. Амалдың орындалу алгоритмі келесідей: **Сервис/Анализ данных/Регрессия/Ок;(**1.7-сурет**).**



1.7-сурет. «Регрессия» капшығын шығару

Кіргізу интервал У; У мәні бар бағанды белгілеңіз (С2:С10;);

Шығару интервал Х; Х мәні бар бағанды белгілеңіз (В2:В10);

Шығару интервал: Нәтижелерді шығару үшін бос ұяшық аймағын белгілеңіз, мысалы, В12.

Excel шешімдерді 1.12–1.14-кестелер түрінде көрсетеді.

1.12-кесте

Регрессиялық статистика

Регрессиялық статистика	
Жиынтық R	0,995873
R-квадрат	0,991762
Калыпты R-квадрат	0,990586
Стандарттық қате	0,370885
Бақылау саны	9

1.13-кесте

Дисперсиялық талдау

	df	*SS*	*MS*	*F*
Регрессия	1	115,926	115,926	842,7576737
Қалдық	7	0,962888889	0,137555556	
Барлығы	8	116,8888889		

df=1 – еркіндік деңгейінің саны;

SS – айырымның квадраттарының сомасын:

1-ші еркіндік деңгей санымен регрессия квадраттарының сомасы SS_1=115,926.

n-2 еркіндік деңгей санымен қалдық квадраттарының сомасы –SS_2=0,963.

n-1 еркіндік деңгей санымен жалпы квадраттар сомасы – 116,889.

MS-дисперсия бағасы:
Регрессия дисперсиясы – $\delta^2_{факт} = SS_1 / 1 = 115,926$;
Қалдық дисперсиясы – $\delta^2_{калд} = SS_2 (n-2) = 0,138$;
F-Фишер критерийі: F=842,75.

1.14-кесте

Регрессия параметрінің теңдеуі

	Коэф-фициент-тер	*Стандарт-тық қате*	*t-ста-тисти-касы*	*P-мәні*	төменгі шекара *95%*	жоғарғы шекара *95%*
1	2	3	4	5	6	7
Y-қиылысу	1,0289	0,2280	4,5135	0,0028	0,4899	1,5679
X 1	1,39	0,0479	29,0303	0,0000	1,2768	1,5032

Регрессия параметрінің теңдеуі осы 1.14-кестеде көрсетілген: 2-ші графа a және b коэффициентінің мәнін көрсетеді: a=1,0289, b=1,39. 3-ші графа стандарттық қателіктер; t_a=0,228 және t_b=0,0479;

4-баған *t*-статистика; I_a=4.5135, I_b=29,0303. 5–6-баған сенімді интервалдар. P=0,95 сенімді ықтималдықпен регрессия параметрлері үшін интервалды бағалар

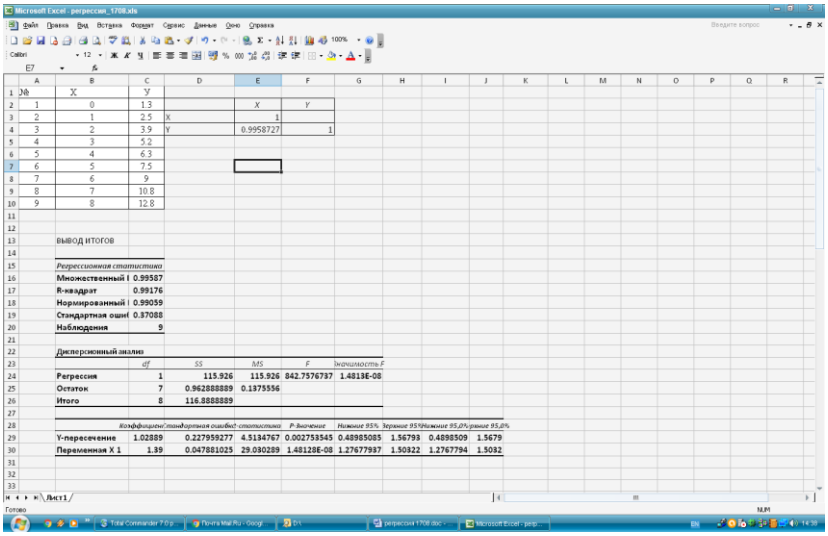
γ_{amin}=0,4899, γ_{amax}=1.5679, γ_{bmin}= 1,2768, γ_{bmax}=1,5032, p=0,95.

Нәтиже бойынша регрессия тендеуін жазып қоялық У=1,0289+1,39*x.

Регрессия параметрлерінің сенімді интервалдары

$$0,4899<a<1,5679, \qquad 1,2768<b<1,5032.$$

Екі әдіспен есептеу кезінде қателіктер дөңгелектеумен байланысты. Excel-де арнаулы пакеттерді қолдану есептеудің дәлдігін қамтамасыз етеді (1.8-сурет).



№	X	У
1	0	1.3
2	1	2.5
3	2	3.9
4	3	5.2
5	4	6.3
6	5	7.5
7	6	9
8	7	10.8
9	8	12.8

	X	Y
X	1	
Y	0.9958727	1

ВЫВОД ИТОГОВ

Регрессионная статистика

Множественный I	0.99587
R-квадрат	0.99176
Нормированный I	0.99059
Стандартная ошиб	0.87088
Наблюдения	9

Дисперсионный анализ

	df	SS	MS	F	Значимость F
Регрессия	1	115.926	115.926	842.7576737	1.4813E-08
Остаток	7	0.962888889	0.1375556		
Итого	8	116.8888889			

	Коэффициенты	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%	Нижние 95,0%	Верхние 95,0%
Y-пересечение	1.02889	0.227959277	4.5134767	0.002753545	0.489850851	1.56793	0.4898509	1.5679
Переменная X 1	1.39	0.047881025	29.030289	1.48128E-08	1.276779371	1.50322	1.2767794	1.5032

1.8-сурет. Регрессиялық талдау

1.20-тапсырма. Өндірістік және өткізу қорларының корреляциялық тәуелділігін есептеу. х – тауарлық қор, млн.тг.; у – өндірістік қор, млн.тг. Көрсетілген мәліметтер бойынша корреляциялық-регрессиялық талдау жасау [9,10,32].

№	1	2	3	4	5	6
Х	9,3	7,5	7	8,3	7,5	6,2
У	1,3	1,3	0,8	0,9	1,4	1,1

1.21-тапсырма. Берілгені: х – тауарлық қор, мың тг.; у – сатылым көлемі, мың тг./тәулік. Көрсетілген мәліметтер бойынша корреляциялық-регрессиялық талдау жасау.

Х	410	395	228	320	290	325
У	12,3	11	5,4	6,2	6,0	8,2

1.22-тапсырма. Берілгені: х – жабдықтандыру сенімділігі, 0<R<1;

у – өндірістік қор мөлшері, млн. тг.

x	0,72	0,65	0,83	0,56	0.51	0,67
y	5,3	4,6	2,3	6,7	7,9	5,7

1.23-тапсырма. Келтірілген мәлімет бойынша **x** жеке табыс пен **y** тамақ шығындарының тәуелділігінің түрі келесідей болады: $\hat{y} = \underset{(1,65)}{6,25} + \underset{(0,143)}{0,775}x$

Жақшада стандартты қателер көрсетілген: $\mathbf{S_a} = 1,65$; $\mathbf{S_b} = 0,143$.

b=0,775 коэффициентінің мәнділігін бағалау және 5 % мәнділік деңгейінде β үшін сенімді интервалын құру.

1.10. Сызықтық регрессия коэффициенттерінің статистикалық маңыздылығы

Ең аз квадраттар тәсілімен есептелген регресиия коэффициенттері – бұл қасиеттері теңдеудегі қалдық мүшенің қасиеттеріне байланысты кездейсоқ шаманың ерекше түрі. Мұны алдымен теория жүзінде, содан соң бақыланатын тәжірибе көмегімен көрсетейік. Жеке алғанда, қалдық мүшеге қатысы бар кейбір нақтылы болжаулардың регрессия коэффициенттерін бақылау үшін қандай мағынасы бар екендігін көреміз [3,9].

Қарастыру барысында біз үнемі x пен y келесі тәуелділікте байланысқан жұптық регрессия теңдеуін жазамыз.

$$y=\alpha+\beta x+u \qquad (1.25)$$

(1.25) және n ішінара бақылау негізінде регрессия теңдеуін бағалаймыз.

$$\hat{y}=a+bx \qquad (1.26)$$

Біз сондай-ақ x – бұл кездейсоқ емес экзогенді айнымалы деп болжаймыз. Басқаша айтқанда, оның мәні барлық бақылауларда алдын-ала берілген деп және зерттелетін тәуелсіздікке еш қатысы жоқ деп ұйғарамыз.

Біріншіден, y шамасы екі құрамдас бөліктен тұратынын байқаймыз. Ол, ықтималдық заңдарымен байланысы жоқ (α+βx) кездейсоқ емес құрамдасынан (α мен β белгісіз болуы мүмкін, дегенмен де бұл тұрақты шамалар) және u кездейсоқ құрамдасынан тұрады.

Осыдан, біз кәдімгі

$$b = \frac{Cov(x, y)}{Var(x)} \tag{1.27}$$

формуласымен есептегенде b-да кездейсоқ құрамдас бар екендігі шығады. Cov(x,y) y-тің мәндеріне, ал y u-дың мәндерінен байланысты.

Егер кездейсоқ құрастырушы n бақылауларында әртүрлі мән қабылдайтын болса, онда y-тің әртүрлі мәндерін аламыз, және сондықтан әртүрлі шамалар Cov(x,y) және b.

Теория жүзінде b-ны кездейсоқ және кездейсоқ емес құрастырушыларға жіктей аламыз. (1.25) қатынасты қолданып сондай-ақ ковариациямен есептеу ережелерін қолданып, табамыз:

$$Cov(x,y) = Cov(x,[\alpha+\beta x+u]) = Cov(x,\alpha) + Cov(x,\beta x) + Cov(x,u)\ . \tag{1.28}$$

3-ші ковариациялық ереже бойынша Cov(x,a) нөлге тең. 2-ші ковариациялық ереже бойынша Cov(x, βx) βCov(x,x)-ке тең. Және де Cov(x,x) бұл Var(x)-пен бірдей. Сондықтан:

$$Cov(x, y) = \beta\ Var(x) + Cov(x, u), \tag{1.29}$$

деп жаза аламыз, және осылайша

$$b = \frac{Cov(x, y)}{Var(x)} = \beta + \frac{Cov(x, u)}{Var(x)}. \tag{1.30}$$

Сонымен, біз кез келген сұрыптау бойынша алынған b регрессия коэффициенті екі қосылғыштың қосындысы ретінде:

b коэффициентінің дәл мәніне тең тұрақты шама;

β тұрақтысынан b коэффициентінің ауытқуына себепші болатын, Cov(x,u)-дан тәуелді кездейсоқ құрастырушымен берілетінін көрсеттік.

Осылайша, *a*-да α-ның дәл мәніне тең болатын тұрақты құрастырушы қосылған u кездейсоқ факторынан тәуелді кездейсоқ құрастырушы бар екендігін көрсетуге болады.

Іс жүзінде біз регрессия коэффициенттерін кездейсоқ құрастырушыларға жіктей алмайтындығымызды ескерткен жөн, өйткені α мен β-ның дәл мәндерін немесе сұрыптаудағы u-дың өзіндік мәндерін білмейміз. Олар бізге белгілі бір пайымдауларда a және b теориялық қасиеттері туралы кейбір ақпарат алуға мүмкіндік беретіндіктен қажет.

1. Бағаланған шешім b, интервал шетіне өседі, яғни гипотез Но-ғы әділеттілікте b-ге тең мүмкінділік жағдайы болады.

Мүмкінділік b аз болғандықтан, келесі шешім жасалады; гипотеза Но регрессияны бағалау шешіміне қарсы келеді деп, және оны аз маңызды жағдай болды дегенше қарастырмау шынайы болып келеді деп есептеледі.

2. Бағаланған шешім интервалында өседі яғни 1- b -ге мүмкіндік тең жағдайы болып жатыр. Бұл жағдайда гипотеза Но қарастырылмай қоймайды.

Гипотезді тексеру кезінде екі типті қателер болуы мүмкін. 1-ші түрдегі қате – нақ нөлдік гипотеза қарастырылмайды, ал 2-ші түрдегі қате рас емес нөлдік гипотеза болса қарастырылмай қоймайды. В артуы 1-ші түрдегі қатеге алып келеді, ал азаюы 2-ші түрдегі қателікке алып келеді.

Бақылау сұрақтары

1. Регрессия коэфиценттерінің екі буыннан тұратындығын дәлелдеңіз.

2. Жорамалды тексеруде пайда болатын I және II түрдегі қателердің мәні неде?

3. *a* регрессия коэффиценті қандай буындарда тұратындығын атаңыз.

1.11. Монте-Карло әдісі бойынша эксперимент

Бұл тәжірибенің не үшін Монте-Карло әдісі бойынша аталатынын ешкім білмейді. Бұл атау кездейсоқ заңдылыққа қатысты болуы мүмкін. Негізгі түсінік ұқсастық арқылы түсіндіріледі. Шошқа трюфельдерді табуға үйретілген деп ұйғарайық. Бұл саңырауқұлақ Франция мен Италияда өсетін жабайы топырақты, айтулы тамақ. Табу қиын болғандықтан саңырауқұлақтар қымбат және оларды табатын жақсы шошқа да қымбат тұрады. Мұндағы мәселе шошқаның қаншалықты трюфельдерді табатынынан тұрады. Ол трюфелдерді уақыт өткен сайын тауып немесе оларды жіберіп қоюы да мүмкін. Шынында қызығушылық танытып сіз жер учаскесін алып, трюфельдерді бірнеше жерге көміп, шошқаны жіберіп оның қаншасын табатынын көруіңізге болады. Бұл эксперимент арқылы іздеудің жемісті дәрежесін тура бағалауға болады [3,4,9].

Мұның регрессиялық талдауға қатынасы қандай? Мұндағы мәселе, біздің ешқашан α және β ақиқат мәндерін біле алмайтындығымызда. Соңдықтан біз бұл әдістің жақсы не жаман баға беретінін айта алмаймыз. Монте-Карло эксперименті тексеру нәтижесін бере алатын жасанды түрде бақыланатын эксперимент. Монте-Карло бойынша қарапайым эксперимент үш бөліктен тұрады.

Біріншіден:

1) α және β нақты мәндері таңдалынады;

2) әрбір бақылауда x мәні таңдалынады;

3) әрбір бақылауда u кездейсоқ факторының мәнін алу үшін кездейсоқ сандар өндірудің қандайда бір үрдісі қолданылады.

Екіншіден, әрбір бақылауда $y = \alpha + \beta x + u$ қатынасы мен α, β, x және u мәндерін қолданатын y мәні өндіріледі.

Үшіншіден, тек көрсетілген түрде алынған y мәндерін және x-ке арналған деректерді ғана қолдана отырып a және b мәндерін α және β үшін жақсы немесе нашар баға бола алатынын көре аласыз, аталған әрекеттер регрессия құру әдісінің жорамалданғанын сезінуге мүмкіндік береді.

Бірінші екі қадамда регрессиялық әдістің қолданылуына дайындық жасалады. Біз модельді толықтай бақылаймыз, жасалған параметрлердің ақиқат мәнін білеміз, себебі өзіміз анықтадық. Үшінші этапта біз өзіміз қойған есептің регрессиялық әдістің көмегімен шығарылуын анықтаймыз, яғни α, β үшін тек қана у және х берілгенін пайдаланып жақсы баға алу мүмкіндігін қарастырамыз. Y-ты алу процесінде кездейсоқ факторды қоссақ мәселе туындайтынын аңғарамыз. Бұл фактор жоқ болған жағдайда барлық нүктелер бақылау мәні бойынша дәл жатар еді және де α, β мәндерін *у* және *х* мәндері бойынша оңай анықтауға болар еді.

α=2 және β=0,5 мұндағы шын тәуелділік түрі:

$$y = 2 + 0{,}5x + u. \qquad (1.31)$$

Қарапайым түрде 20 бақылау бар және х 1-ден 20-ға дейінгі мәнді қолданады десек, кездейсоқ қалдықтық құрамдас үшін үлестіріліп қосылған нөлдік орташа және бірлік дисперсиялы кездейсоқ санды қолданамыз. Бізге 20 мәнді қабылдау керек. Оларды *rn1,…rn20* деп белгілейміз. Кездейсоқ мүше *u1* бірінші бақылауда *rn1-ге және т.б.*(1.15-кесте).

1.15-кесте

x	*u`*	*y*	*x*	*u`*	**y**
1	-0.59	1.91	11	1.59	9.09
2	-0.24	2.76	12	-0.92	7.08
3	-0.83	2.67	13	-0.71	7.79
4	0.03	4.03	14	-0.25	8.75
5	-0.38	4.12	15	1.69	11.19
6	-2.19	2.81	16	0.15	10.15
7	1.03	6.53	17	0.02	10.52
8	0.24	6.24	18	-0.11	10.89
9	2.55	9.03	19	-0.91	10.59
10	-0.13	6.87	20	1.42	13.42

Әр бақылауда *x* және *u* мәндерін біліп, (1.31) теңдікті қолданып у мәнін таба аламыз. Енді у-тің х-тен регрессиялық тәуелділігін бағалағанда:

$$\hat{y}=1{,}63+0{,}54x. \qquad (1.32)$$

Бұл жағдайды бағалағанда *a* (1,63) α-мен (2,00) салыстырғанда аз мәнді, ал *b мәні* β мәнінен жоғарырақ (0,54 пен 0,50). Кездейсоқ мүшелердің бірлескен ықпалымен 20 бақылауда айырмашылық келтірілген. Сірә, регрессиялық әдістің сапасын бағалағанда тәжірибенің бір түрі жеткілікті. Ол риза қыларлықтай жақсы нәтиже берді, бірақ бұл бақытты кездейсоқ жағдай болуы мумкін. Мұны әрі қарай тексеру үшін ақиқат теңдікті және х-тің сол ақиқат мәндерін алып тәжірибені қайталаймыз, бірақ кездейсоқ сандарды жаңадан теру қалған мүше үшін сол бөлуден алынған (нөлдік орташа және бірлік дисперсиясы). Бұл мәндерді және х-тің мәндерін қолданып, у-тің жаңа теруін аламыз. Мына нәтиже у пен х-тің жаңа мәндерінің арасында регрессияны бақылаудан шықты:

$$\hat{y}=2{,}52+{,}48x. \qquad (1.33)$$

Сондай-ақ екінші тәжірибеде жемісті болды. Енді а α-дан көп болса, ал *b,* β мәнінен бірнешеге аз болды. 1.16-кестеде кездейсоқ сандарды әр жағдайда әртүрлі терулерді қолданып, 10 рет қайталау тәжірибесімен а және b бағалары келтірілген. Бір жағдайда бағалардың төмен мәндерді қабылдауына, басқаларында жоғары болуына қарамастан а және b тұтастай мәндері сәйкесінше 2,00 және 0,50-ге тең α және β мәндерінің айналасына топтасады. Мұнда нашар бағалардан гөрі жақсы бағалар көп алынды. Мысалға, тәжірибені өте көп қайталау кезінде b мәнін бекітіп, жиіліктер кестесі мен аппроксимациялық функцияның тығыздық ықтималдығын алуға болады.

1.16-кесте

Тәжірибе	*a*	*b*	Тәжірибе	*a*	*b*
1	1.63	0.54	6	1.81	0.51
2	2.52	0.48	7	1.72	0.56
3	2.13	0.45	8	3.18	0.41
4	2.14	0.50	9	1.26	0.58
5	1.71	0.56	10	1.94	0.52

Жоғарыда айтылғандай, регрессиялық коэффициенттер мен ақиқат мәндер арасындағы *u* кездейсоқ мүшесімен келтірілген. Осыдан келе, кездейсоқ элемент көп болған сайын бағалардың дәлдігі азаяды. Бұл нәтиже Монте-Карло әдісінің екінші сериялы тәжірибесімен суреттелетін болады, ол біріншісімен байланысты. Біз бұрынғыдай α және β, x мәндерін, кездейсоқ сандардың қайнар көздерін кездейсоқ мүшені генерациялау үшін қолданамыз, бірақ енді кездейсоқ мүшенің мәндерін әр бақылауда аламыз.

Соңғысын *u`i (i*=1,2,…,*n)*, арқылы білдіреміз, мұндағы мән екі еселенген кездейсоқ санға тең: *u`1=2rn1,…,u`20=2rn20*. Фактіге негізделген дәл сол кездейсоқ сандарды қолданамыз, бірақ бұл ретте олардың мәндерін екі еселейміз. Енді 1.15-кестенің негізінде 1.17-кестені есептейміз. Әрмен қарай y және x-тің арасындағы регрессияны бағалауда теңдік аламыз:

$$\hat{y}=1{,}26+0{,}58x. \qquad (1.34)$$

Бұл теңдік (1.32) теңдікке қарағанда анағұрлым дәлсіз болады.

1.17-кесте

X	*u`*	*y*	*x*	*u`*	*y*
1	-1,18	1,32	11	3,18	10,68
2	-0,48	2,52	12	-1,84	6,16
3	-1,66	1,84	13	-1,42	7,08
4	0,06	3,94	14	-0,50	8,5
5	-0,76	3,74	15	3,38	12,8
6	-4,38	0,62	16	0,30	10,3
7	2,06	7,56	17	0,04	10,54
8	0,48	6,48	18	-0,22	10,78
9	5,10	11,56	19	-1,82	9,68
10	-0,26	6,74	20	2,84	14,84

1.18-кесте

Тәжірибе	*a*	*b*	Тәжірибе	*a*	*b*
1	1.26	0.58	6	1.61	0.52
2	3.05	0.45	7	1.44	0.63
3	3.26	0.39	8	4.37	0.33
4	3.28	0.50	9	0.52	0.65
5	1.42	0.61	10	1.88	0.55

1.18-кестеде u`=2rn болғандағы 10-тәжірибенің нәтижесі келтірілген. Біз бұны II сериялы тәжірибе деп, ал алғашқы тәжірибе сериясын – 1-тәжірибелі сериялы дейміз. Оның нәтижесі 1.16-кестеде көрсетілген. 1.16 және 1.18-кестелерді салыстырғанда, a және b мәндері 1.16-кестеде тұрақты емес, дегенмен оларда баяғыдай мәндердің бағасының төмендеуі немесе жоғарылауының жүйелік тенденциясы жоқ.

Бөлшектік зерттеулер бізге маңызды ерекшелікті табуға мүмкіндік береді. I серияда b мәні 1-тәжірибе бойынша 0,54-ке тең болды, бағаның жоғарылауы 0,04 құрады. II серия b мәні 1-тәжірибе бойынша 0,58-ге тең болды, ал жоғарылау 0,08-ді құрады, яғни бұл бұрынғыдан екі есеге көп болды. Дәл осылай 9-тәжірибеде де қайталанады, сонымен бірге a-регрессиялық коэффициенті үшін әр тәжірибеде қайталанады. Әр бақылаудағы кездейсоқ мүшенің екі еселенуі регрессиялық коэффициенттегі қателіктің мәнінің екі еселенуіне әкеледі.

Бұл нәтиже әрі қарай b-ның бұзылуынан сәйкесінше $b = \beta + \frac{Cov(x,u)}{Var(x)}$ теңдікпен жалғасады. I сериядағы b-дағы кездейсоқ қате *Cov (x, u)/Var (x)* түрінде беріледі. II серияда ол *Cov (x, u')/Var (x)*, және

$$\frac{Cov(x,u`)}{Var(x)} = \frac{Cov(x,2u)}{Var(x)} = 2\frac{Cov(x,u)}{Var(x)}. \qquad (1.35)$$

1.12. Гаусс-Марковтың 4-шарты

Регрессия коэффициенттерінің қасиеті кездейсоқ құрастырушы қасиетіне тәуелді. Ең кіші квадраттар әдісіне негізделген регрессиялық талдау мүмкін болатын нәтижелердің ішінен ең жақсы нәтижені беруі үшін кездейсоқ мүше төрт шартты қанағаттандыруы тиіс [3,9,12]:

1-шарт. Кездейсоқ шаманың математикалық үміті нөлге тең, $E(u)=0$.

Кездейсоқ мүше u ешқандай бағытқа ауытқымауы тиіс.

2-шарт. U кездейсоқ мүше дисперсиясы барлық бақылаулар үшін тұрақты, яғни, $D(u)=\sigma^2$. σ^2 шамасы алдын ала белгісіз болады. Егер $D(u)$ дисперсиясы әрбір бақылауда өзгеріп отырса, онда Гаусс-Марков шарты бұзылады.

3-шарт. Кездейсоқ мүшелер барлық бақылауларда бір-бірінен тәуелсіз болуы тиіс. $E(u_i)=E(u_j)=0$ или $E(u_iu_j)=0$. $(i\neq j)$.

Егер бұл шарт орындалмаса ең кіші квадраттар әдісімен бағаланған регрессия тиімді емес нәтижелер бермейді.

4-шарт. Кездейсоқ мүше түсіндіруші айнымалыға тәуелсіз түрде үлестірілген болуы тиіс. Егер x кездейсоқ емес шама болса. Онда т келесідей болуы тиіс: $E(x_iu_i)=0$, $1\le i\le n$.

Бұл төрт шартқа қосымша ретінде кездейсоқ мүше үшін ықтималдықтардың үлестірілуі қалыпты түрде болуы тиіс: $u_i\sim N(0,\sigma^2)$, $1\le i\le n$.

Гаусс-Марков шартына сүйене отырып, a және b коэффициенттерінің келесі қасиеттерін дәлелдеуге болады.

1. a, b шамалары α, β коэффициенттерінің жұп сызықтық регрессия теңдеуіне сәйкес ығыстырылмаған бағалары болып табылады, яғни $E(a)=\alpha$, $E(b)=\beta$.

2. a, b шамалары α, β коэффициенттерінің тиімді бағалары болып табылады.

Регрессия коэффициенттерінің жылжымайтындығы (1.30) теңдеу негізінде, егер Гаусс-Марковтың 4-ші шарты орындалса

$$E\{b\} = E\{\beta + \frac{Cov(x,y)}{Var(x)}\} = \beta + E\{\frac{Cov(x,y)}{Var(x)}\}, \qquad (1.36)$$

b-ның жылжымайтын бағасы болатындығын көрсетуге болады; өйткені β-тұрақты. Егер біз Гаусс-Марковтың 4-ші басым формасын қабылдап және x-кездейсоқ емес шама деп болжасақ, біз Var(x)-ті u-дың белгілі тұрақтысы деп санай аламыз, және осылайша,

$$E\{b\} = \beta + \frac{1}{Var(x)} E\{Cov(x,u)\}. \qquad (1.37)$$

Одан әрі, егер x-кездейсоқ емес шама болса, онда E{Cov(x,u)}=0 және

$$E\{b\} = \beta. \qquad (1.38)$$

Осылайша, b β-ның жылжымайтын бағасы.

n бақылаудағы кездейсоқ факторлар бірін-бірі тура «өшіргенде», бұл тек қана кездейсоқ тұспа-тұс келген жағдайда ғана болады, b β-дан әрбір нақты тәжірибелерде ерекшеленеді. Бірақ та (1.38) қатынасты ескере отырып бағаны жоғарылататын немесе төмендететін жүйелік қате болмайды. Ол а коэффициенті үшін де әділ.

Регрессия коэффициенттерінің дәлдігі

Енді *a* және *b* бағаларының теориялық дисперсияларын қарастырайық. Олар келесі амалдармен беріледі.

$$pop.\mathrm{var}(a) = \frac{\sigma_u^2}{n}\{1 + \frac{\bar{x}^2}{Var(x)}\} \quad \text{және} \quad pop.\mathrm{var}(b) = \frac{\sigma_u^2}{nVar(x)}. \qquad (1.39)$$

(1.38) теңдеуден үш анық қорытынды жасауға болады.

Біріншіден, a және b дисперсиялары σ^2 қалдық мүшесі дисперсиясына тура пропорционал.

Кездейсоқтық факторы көп болған сайын, басқа да тең шарттарда бағалар төмен болады. Бұл Монте-Карло әдісі бойынша тәжірибелерде көрсетілген. II сериядағы бағалар I сериядан қарағанда көбірек дәл емес болды, бұл біздің әрбір сұрыптауда кездейсоқ мүшені екі еселегеннен пайда болды. u-

ды екі еселеп, біз оның стандартты ауытқуын екі еселедік және осылайша, а және b-ның стандартты ауытқуларын екі еселедік. Екіншіден, сұрыптаулар саны көп болған сайын, бағаның дисперсиясы аз болады. Мұның белгілі бір мағынасы бар. Сізде ақпарат көп болған сайын, сіздің бағаларыңыз да дәл болуы ықтимал. Үшіншіден, x дисперсиясы көп болған сайын регрессия коэффициенттерінің дисперсиясы аз болады. Мұның себебі неде? (1) Регрессия коэффициенттерін y-тің бақыланатын өзгерістері x-тің бақыланатын өзгерістері салдарынан болады деген болжаулар негізінде есептеледі, бірақ (2)-де іс жүзінде олар тек x-тің өзгерулерімен, бір жағынан u вариацияларымен өзгерулерінен болған. x дисперсиясы аз болған сайын, y ауытқуын анықтағандағы кездейсоқтық факторының салыстырмалы әсері аз болуы ықтимал. Және регрессиялық талдау дұрыс емес болып шығу ықтималдығы көп. Шынында, (1.39) теңдеуден көрінетіндей абсолютті емес, σ^2 және Var(x) салыстырмалы шамалары маңызды.

Іс жүзінде біз *a* немесе *b* теориялық дисперсияларын есептей алмаймыз, өйткені σ^2 белгісіз, бірақ қалдықтар негізінде σ^2 бағасын ала аламыз. Регрессия сызығына қатысты шашырау y=α+βx сызығына қатысты u белгісіз шашырауын бейнелейтіні анық, бірақ жалпы қалдық және кездейсоқ мүше кез келген бақылауда бір-біріне тең емес. Демек, біз өлшей алатын Var(e) қалдықтардың сұрыптау дисперсиясы, біз ала алмайтын σ^2 бағасы үшін қолданыла алады.

Ары қарай бару үшін, өзіңізге келесі сұрақты қойыңыз: x және y бойынша бақылау сұрыптаулары болатын қандай түзу нүктелерге жақын болады: y=α+βx дәл түзуі немесе $\hat{y}=a+bx$ регрессия сызығы? Жауап былайша болады: регрессия сызығы, өйткені анықтама бойынша ол оның арақашықтықтар квадратының қосындысын және бақылаулар мәндерін ең кішісін алуға әкеліп соғады. Демек, оның қалдықтар шашырауы, u мәндер шашырауынан кем, және Var(e) –да σ_u^2 бағасын төмендету тенденциясы бар. Шынында да, егер бір ғана тәуелсіз

айнымалы бар болса, Var(e) математикалық үміті [(n-2)/n] σ_u^2 тең. Бірақ, егер s^2_u-ды

$$S_u^2 = \frac{n}{n-2} Var(e) \qquad (1.40)$$

деп анықтасақ, онда σ_u^2 σ_u^2-дың жылжымайтын бағасы екендігі шығады.

(1.39) және (1.40) қолданып, a және b үшін теориялық дисперсиялар бағасын олардың стандартты ауытқу бағаларынан түбір астына алғанда да табуға болады. Регрессия коэффициентінің шамасы «ықтималдықтар тығыздығы функциясының стандартты ауытқу бағасы» терминінің орнына регрессия коэффициентінің «стандартты қатесі» терминін «с.қ.» қысқа түрде белгілейміз. Осылайша, жұптық регрессиялық талдау үшін:

$$\text{с.қ.}(a)=\sqrt{\frac{s_u^2}{n}\{1+\frac{\bar{x}^2}{Var(x)}\}} \text{ және } \text{с.қ.}(b)=\sqrt{\{\frac{s_u^2}{nVar(x)}\}}. \qquad (1.41)$$

Алынған қатынастар Монте-Карло әдісі бойынша тәжірибелермен көрсетілген I серияда, u рөлдік орта және (δ^2_u=1) бірлік дисперсиясы бар бас жиынтықтан алынған кездейсоқ сандар негізінде анықталды, ал x 1-ден 20-ға дейінгі сандар жиынтығы болатын 33,25-ке тең болатын Var(x) оңай есептеуге болады

$$pop.\text{var}(a) = \frac{1}{20}\{1+\frac{10{,}5^2}{33{,}25}\} = 0{,}2158 \qquad (1.42)$$

және

$$pop.\text{var}(b) = \frac{1}{20*33{,}25} = 0{,}001504. \qquad (1.43)$$

Осылайша, b үшін дәл стандартты ауытқу мәні $\sqrt{0.001504}$ =0.039 тең. (1.25) теңдеуді қолданып стандартты қатені есептегенде I серияның 10-тәжірибесінде қандай нәтижелер алынған; 10-тәжірибе үшін бұл есептеулердің нәтижелері 1.19-кестеде берілген. Көріп отырғанымыздай, көпшілік бағалар жеткілікті жақсы.

1.19-кесте

Тәжірибе	С.қ. (b)	Тәжірибе	С.қ. (b)
1	0,043	6	0,044
2	0,041	7	0,039
3	0,038	8	0,040
4	0,035	9	0,033
5	0,027	10	0,033

Тағы бір негізгі кезеңді айтып кету керек. Стандарттық қате тек регрессия коэффициентінің дәлдік дәрежесінің жалпы бағасын ғана береді. Ол, алынған баға таралудың орташасында болатындығы туралы ақпарат бермейді және осылайша дәл немесе таралудың «соңында» бола ма бұл салыстырмалы түрде дәл емес.

Кездейсоқ мүшенің дисперсиясы көп болған сайын қалдықтардың сұрыптау дисперсиясы көп болатындығы анық, және регрессия теңдеулеріндегі коэффициенттердің стандартты қателері, ол алынған коэффициенттердің дәл еместігін жоғары ықтималдықпен қорытындылауға мүмкіндік береді. Бірақ бұл тек қана ықтималдық. Белгілі бір сұрыптауда әртүрлі бақылауларда кездейсоқ фактордың әсерлері өзара өшірілуі мүмкін және нәтижесінде регрессия коэффициенттері дәл болады. Мәселе, былайша айтқанда, мұның болатындығын немесе болмайтындығын бекітуге болмайтындығында.

Гаусс-Марков теоремасы.

Бізде y=α+βx+u теңдеумен берілген тәуелділік бар делік, және назарды β үшін бағалауларға аударайық. Регрессиялық талдаумен таныс емес адам, бақылаулар сұрыптауы үшін шашырау диаграммасын көріп бірінші және соңғы бақылауларды біріктіру жолы арқылы және биіктіктің өсуін олардың арасындағы горизонталды кесіндіге бөлу арқылы көлбеу бұрыш тангенсін бағалауды алуға ұмтылуы мүмкін. Бұл жағдайда b-ның бағасы келесі жолмен анықталатын болады:

$$b = \frac{y_n - y_1}{x_n - x_1}. \qquad (1.44)$$

Бұл бағаның қасиеттері қандай? Алдымен ол жылжытылған ба соны зерттейік. $y=\alpha+\beta x+u$ теңдеуді қолданып бірінші және соңғы бақылауларға қатысты аламыз.

$$y_1=\alpha+\beta x_1+u_1 \qquad (1.45)$$

$$y_n=\alpha+\beta x_n+u_n \qquad (1.46)$$

Демек,

$$b=\frac{\beta x_n+u_n-\beta x_1-u_1}{x_n-x_1}=\beta+\frac{u_n-u_1}{x_n-x_1}. \qquad (1.47)$$

Осылайша, біз, «қарапайым» бағаны екі құрастырушыға жіктедік: дәл мән және қалдық мүше. $E(u)=0$ деп алып u_1 және u_n математикалық күтулері нөлге теңдігін аламыз, бірақ (1.47) теңдеудегі қалдық мүшенің математикалық үміті де нөлге тең. Осылайша, бұл бағаның соншама «қарапайым» екеніне қарамастан, ол жылжытылған болып табылады.

Бұл өзі де, ең кіші квадраттар әдісімен алынған бағамен қатар жалғыз баға емес, жылжытылғандық қасиеті бар. Сіз осы типті тағы бір бағаны кез келген алынған бақылауларды біріктіру арқылы ала аласыз, ал егер сіз одан да «қарапайым» процедураларды қарастыратын болсаңыз, онда бұл жерде шексіз мүмкіндіктер ашылады.

1.6-мысал. Айталық, әйтеуір бір елде 5 жылдағы *y* тамаққа кеткен шығын мен *x* жеке табыс (шартты бірлік) кестесі берілген болсын (1.20-кесте). Қажет:

1) y және x регрессиялық тәуелділікті тұрғызу.

2) Реттеу сапасын бағалау.

3) Регрессия коэффициенттерінің стандартты қателерін есептеу.

Шешуі:

1.20-кесте

Жыл	x	y	x^2	x*y	$\hat{y}$	$(y-\hat{y})^2$	$(\hat{y}-\bar{y})^2$	ℓ^2	$(x-\bar{x})^2$
2000	2	9	4	18	7,8	25	38,44	1,44	64
2001	6	10	36	60	10,9	16	9,61	0,81	16
2002	10	12	100	120	14	4	0	4	0
2003	14	19	196	266	17,1	25	9,61	3,61	16
2004	18	20	324	360	20,02	36	38,44	0,04	64
Бар-лығы	50	70	660	824	70	106	96,1	9,9	160
Орташа	$\bar{x}=\frac{1}{5}\sum_1^5 x_i$ =10	$\bar{y}=\frac{1}{5}\sum y_i$ =14	$\overline{x^2}=132$	$\overline{x\cdot y}$ =164.8	$\bar{y}=14$	Var(y) =21.2	Var($\hat{y}$) =19.22	Var(e) =1.98	Var(x) =32

1) ЕККӘ бойынша

$$b=\frac{\operatorname{cov}(x,y)}{\operatorname{var}(x)}=\frac{\overline{x*y}-\bar{x}*\bar{y}}{\bar{x}^2-(\bar{x})^2}=\frac{164.8-140}{132-100}=\frac{24.8}{32}=0.775;$$

b=0,775, $a=\bar{y}-b\bar{x}=14-0.775*10=6.25,$

регрессия теңдеуі: $\hat{y}=a+bx, \hat{y}=6.25+0.775x.$

b коэффициенті табыс 1 шартты бірлікке өскенде тамақ шығындары орта есеппен 0,775 шартты бірлікке артатынын көрсетеді.

2) Реттеу сапасын бағалаймыз, яғни детерминация коэффициентін есептейміз: $R^2=\frac{\operatorname{var}(\hat{y})}{\operatorname{var}(y)}$, ол үшін есептеу кестесін жалғастырамыз $y_i-\bar{y}$, $\hat{y}_i=a+bx_i$, $e_i=y_i-\hat{y},$

$$\operatorname{var}(\hat{y})=\frac{1}{n}\sum_1^n(\hat{y}_i-\bar{y})^2,\ \operatorname{var}(y)=\frac{1}{n}\sum(y_i-\bar{y})^2,\ R^2=\frac{19.22}{21.2}=0.907.$$

у тәуелді айнымалының 90,7% вариациясы регрессия теңдеуімен түсіндіріледі.

3) Стандартты қателерді S_a және S_b есептейміз:

Келесі есептеу жүргіземіз: $\operatorname{var}(x)=\frac{1}{n}\sum_1^5(x_i-\bar{x})^2=32,$

$Var(e)=\frac{1}{n}\sum_1^5 e_i^2=1.98$ n = 5.

Қалдық дисперсия:

$$\mathbf{S^2}=\frac{n}{n-2}\operatorname{var}(e)=\frac{5}{3}*1.98=3.3.$$

Регрессияның стандартты қатесі: $S=\sqrt{S^2}=\sqrt{3.3}=1.82$.

Регрессия коэффициенттерінің стандартты қателері:

$$S_a=S\sqrt{\frac{\overline{x^2}}{n*\operatorname{var}(x)}}=1.82*\sqrt{\frac{132}{5*32}}=1.65;$$

$$S_b=\frac{S}{\sqrt{n*\operatorname{var}(x)}}=\frac{1.82}{\sqrt{5*32}}=0.143.$$

1.24-тапсырма. Бос мүшесі жоқ регрессия моделі үшін *y* тамаққа кеткен шығындардың *x* жеке табыстан тәуелділігін тұрғызыңыз және регрессия коэффициенттерінің стандартты қателерін есептеңіз.

Кіріс мәліметтер 1.21-кестеде келтірілген:

1.21-кесте

Жыл	x	y
2007	2	9
2008	6	10
2009	10	12
2010	14	19
2011	18	20

Бақылау сұрақтары

1. Гаусс-Марковтың 4-шарты не үшін қолданылады?
2. Гаусс-Марковтың барлық 4 шартын айтып шығыңыз және олардың ерекшеліктері неде?
3. *a* және b коэффициенттерінің қасиеттерін айтып шығыңыз.
4. Жылжытылғандық дегеніміз не?
5. Гаусс-Марков теоремасын құрастырып айтыңыз?
6. Гаусс-Марковтың төрт шарты не үшін қолданылады?
7. Гаусс-Марковтың төрт шартын атаңыз және олардың ерекшелігін көрсетіңіз?
8. *a* және b коэффициенттерінің қасиеттерін атап беріңіз.

1.13. Сызықтық емес экономикалық модельдер

Егер экономикалық құбылыстар арасында сызықтық емес қатынастар болса, онда олар тиісті сызықтық емес функциялар көмегімен өрнектеледі.

Сызықтық емес регрессиялардың екі класын ажыратады [3,9,12,16]:

1. Талдауға енгізілген түсіндіруші айнымалыларға қатысты сызықтық емес, бірақ бағаланылатын параметрлер бойынша сызықты регрессиялар, мысалы,

– әртүрлі дәрежедегі полиномдар – $\hat{y}_x = a + bx + cx^2$, $\hat{y}_x = a + bx + cx^2 + dx^3$;

– тең қабырғалы гипербола – $\hat{y}_x = a + b / x$;

– жартылай логарифмді функция – $\hat{y}_x = a + b * \ln x.$

2. Бағаланылатын параметрлер бойынша сызықты емес регрессиялар, мысалы,

– дәрежелі – $\hat{y}_x = a + x^b$;

– көрсеткіштік – $\hat{y}_x = a + b^x$;

– экспоненциалды – $\hat{y}_x = e^{a+bx}$.

Енгізілген айнымалылар бойынша сызықтық емес регрессиялар айнымалыларды қарапайым алмастыру тәсілімен сызықтық түрге келтіріледі, ал параметрлердің ары қарай бағалануы ең кіші квадраттар әдісі көмегімен жүзеге асырылады.

Кейбір функцияларды қарастырайық.

Екінші дәрежелі $\hat{y}_x = a + bx + cx^2$ параболасы келесі $x = x_1$, $x^2 = x_2$ алмастыру көмегімен сызықтық түрге келтіріледі. Нәтижесінде екі факторлы $\hat{y}_x = a + bx_1 + cx_2$ теңдеуге келеміз, оның параметрлерін ең кіші квадраттар әдісі

көмегімен бағалау арқылы келесі нормалды теңдеулер жүйесін аламыз:

$$\begin{cases} an + b\sum x_1 + c\sum x_2 = \sum y; \\ a\sum x_1 + b\sum x_1^2 + c\sum x_1 x_2 = \sum x_1 y; \\ a\sum x_2 + b\sum x_1 x_2 + c\sum x_2^2 = \sum x_2 y. \end{cases}$$

Ал айнымалыларды қайта кері алмастырғаннан кейін келесі өрнекті аламыз:

$$\begin{cases} an + b\sum x + c\sum x^2 = \sum y; \\ a\sum x + b\sum x^2 + c\sum x^3 = \sum xy; \\ a\sum x^2 + b\sum x^3 + c\sum x^4 = \sum x^2 y. \end{cases}$$

Екінші дәрежелі парабола фактордың мәндерінің белгілі интервалы үшін қарастырылып отырған мәндердің байланыстарының сипаты өзгертілетін жағдайларда қолданылады: тура байланыс кері байланысқа өзгертілген жағдайда немесе керісінше.

Тең қабырғалы гипербола $\hat{y}_x = a + b/x$ шикізаттың, материалдардың, отынның шығарылған өнім көлеміне байланысты, тауардың айналым уақытының тауарайналым көрсеткішіне байланысты, жалақының өсу пайызының жұмыссыздық деңгейіне байланысты, азық-түлік емес тауар шығындарының табысқа немесе шығындардың жалпы сомасына байланысты үлесті шығындарының байланысының сипаттамасы үшін қолданылуы мүмкін. Гипербола қарапайым алмастыру тәсілімен сызықтық түрге келтіріледі: $z = 1/x$. Ең кіші квадраттар әдісі көмегімен бағалау жүргізген соң келесі теңдеулер жүйесін аламыз:

$$\begin{cases} an + b\sum \frac{1}{x} = \sum y; \\ a\sum \frac{1}{x} + b\sum \frac{1}{x^2} = \sum \frac{1}{x} y. \end{cases}$$

Дәл осылайша, келесі тәуелділіктер $\hat{y}_x = a + b\ln x$, $\hat{y}_x = a + b\sqrt{x}$ сызықтық түрге келтіріледі.

Бұл жағдай бағаланылатын параметрлер бойынша сызықты емес регрессияларда өзгеше өтеді, олар екіге бөлінеді: іші сызықты сызықты емес модельдер (тиісті жаңартулар көмегімен сызықтық түрге келтіріледі, мысалы, логарифмдеу) және іші сызықты емес сызықты емес модельдер (сызықтық түрге келтірілмейді).

Іші сызықты модельдерге, мысалы, дәрежелі – $\hat{y}_x = a + x^b$, көрсеткіштік – $\hat{y}_x = a + b^x$, экспоненциалды – $\hat{y}_x = e^{a+bx}$, логистикалық – $\hat{y}_x = \frac{a}{1+be^{-cx}}$, кері – $\hat{y}_x = \frac{1}{a+bx}$ функциялар жатады.

Іші сызықты емес модельдерге, мысалы, келесі модельді жатқызуға болады: $\hat{y}_x = a + bx^c$, $\hat{y}_x = a + \left(1 - \frac{1}{1-x^b}\right)$.

Сызықты емес модельдердің арасында $\hat{y}_x = a + x^b\varepsilon$ дәрежелік функция жиі қолданылады, ол сызықтық түрге логарифмдеу арқылы келтіріледі:

$$\ln y = \ln(ax^b\varepsilon);$$

$$\ln y = \ln a + b\ln x + \ln\varepsilon;$$

$$Y = A + bX + E.$$

Мұнда $Y = \ell ny$, $X = \ln x$, $A = \ln a$, $E = \ln\varepsilon$. Яғни біз ЕКӘ жаңартылған мәліметтер үшін қолданамыз:

$$\begin{cases} An + b\sum X = \sum Y \\ A\sum X + b\sum X^2 = \sum XY \end{cases}$$

Сосын потенциалдау арқылы жанама теңдеуді аламыз.

Дәрежелік функцияның жиі қолданылуы b параметрінің нақты экономикалық түсіндірілуіне байланысты – ол икемділік коэффициенті болып табылады (икемділік коэффициенті фактор 1 %-ға өзгергенде нәтиже қанша пайызға өзгеретінін көрсетеді).

Икемділік коэффициентін есептеу формуласының жалпы түрі келесідей:

$$Э = f'(x)\frac{x}{y}. \qquad (1.48)$$

Басқа функциялар үшін икемділік коэффициенті тұрақты шама болмағандықтан, *x* факторының тиісті мәніне тәуелді болғандықтан, көбінесе икемділік коэффициентінің орташа мәні:

$$\overline{Э} = f'(x)\frac{\overline{x}}{\overline{y}}. \qquad (1.49)$$

Регрессияның жиі қолданылатын түрлері үшін икемділік коэффициентінің орташа мәнін есептеу формулаларын келтірейік:

1.22-кесте

Функция түрі	Бірінші туынды	Икемділіктің орташа коэффициенті, $\overline{Э}$
1	2	3
$y = a + bx + \varepsilon$	b	$\frac{b\overline{x}}{a + b\overline{x}}$
$y = a + bx + cx^2 + \varepsilon$	$b + 2cx$	$\frac{(b + 2c * \overline{x})\overline{x}}{a + b\overline{x} + c * \overline{x}^2}$
$y = a + b / x + \varepsilon$	$\frac{b}{x^2}$	$\frac{b}{a\overline{x} + b}$
$y = ax^b\varepsilon$	abx^{b-1}	b
$y = ab^x\varepsilon$	$a * \ln b * x^x$	$\overline{x}\ln b$
$y = a + b\ln x + c$	$\frac{b}{x}$	$\frac{b}{a + b\ln\overline{x}}$
$y = \frac{a}{1 + be^{-cx+\varepsilon}}$	$\frac{abce^{-cx}}{(1 + be^{-cx})^2}$	$\frac{b * c * \overline{x}}{b + e^{\overline{cx}}}$
$y = \frac{1}{a + bx + \varepsilon}$	$\frac{b}{(a + bx)^2}$	$\frac{b\overline{x}}{a + b\overline{x}}$

Икемділік коэффициентінің мәнін есептеудің қажеттігі жоқ жағдайлар да кездеседі. Бұл қарастырылып отырылған мәндер үшін пайызбен өзгеруін анықтау қажетсіз болған жағдайда кездеседі.

Сызықты емес регрессия теңдеуі, сызықтық тәуелділік жағдайындағыдай байланыс тығыздығының көрсеткішімен толықтырылады. Берілген бұл корреляция индексі:

$$\rho_{xy} = \sqrt{1 - \frac{\sigma^2_{калд}}{\sigma^2_y}},$$

мұндағы $\sigma^2_y = \frac{1}{n}\sum (y - \bar{y})^2 -$ y нәтижелік мәнінің жалпы дисперсиясы, $\sigma^2_{калд} = \frac{1}{n}\sum (y - \hat{y}_x)^2$ – қалдық дисперсия.

Берілген көрсеткіш $0 \le \rho_{xy} \le 1$ аралығында жатады. Корреляция индексінің мәні бірге жақындаған сайын, қарастырылып отырған мәндердің байланысы тығызырақ, ал регрессия теңдеуі сенімдірек болады. Корреляция индексінің квадратының атауы детерминация индексі және ол нәтижелік мәнінің жалпы дисперсиясында регрессиямен түсіндірілетін y нәтижелік мәнінің дисперсиясының үлесін сипаттайды:

$$\rho^2_{xy} = 1 - \frac{\sigma^2_{калд}}{\sigma^2_y} = \frac{\sigma^2_{факт}}{\sigma^2_y}. \qquad (1.50)$$

Яғни оның мәні сызықты регрессиядағыдай;

$$\sigma^2_{факт} = \frac{1}{n}\sum (\hat{y}_x - \bar{y})^2.$$

Сызықты функцияны қолдану мүмкіндігін негіздеу үшін ρ^2_{xy} детерминации индексін r^2_{xy} детерминация коэффициентімен салыстыруға болады. Мұнда регрессия сызығының қисығы үлкен болған сайын, r^2_{xy} мәні ρ^2_{xy} мәнінен аз болады. Ал бұл екі көрсеткіштің жақындығы регрессия теңдеуінің формасын күрделендіру қажеттігі жоқ екенін көрсетеді және бұл жағдайда сызықтық функцияны қолдануға болады.

Детерминация индексі Фишердің *F*-критерийі регрессияның теңдеуінің мәнділігін тексеру үшін қолданылады:

$$F = \frac{\rho_{xy}^2}{1-\rho_{xy}^2} \bullet \frac{n-m-1}{m}, \qquad (1.51)$$

мұндағы ρ_{xy}^2 – детерминация индексі, n – зерттеулер саны, $m-x$ айнымалысы кезінде параметрлер саны. F-критерийдің (1.51) нақты мәні мәнділік деңгейі α және еркіндік дәрежесі $k_2 = n-m-1$(квадраттардың қалдық сомасы үшін) және $k_1 = m$ (квадраттардың факторлы сомасы үшін) кестелік мәнімен салыстырылады.

Регрессияның сызықтық теңдеуінің сапасы туралы орташа аппроксимация қателігі бойынша қарастыруға болады.

1.7-мысал. Отбасылардың сегіз тобына өткізілген сұраныстың мәліметтері бойынша отбасы табысының деңгейіне сәйкес халықтың тамақ өнімдерінің шығындарының байланысы бойынша мәліметтер берілген. Бұл белгілер арасындағы байланыстың сипаты сызықты емес деп ұйғарып, келесі теңдеудің параметрлерін табамыз:

$$y = a + b\ln x + \varepsilon .$$

1.23-кесте

Бастапқы мәліметтер

Халықтың тамақ өнімдерінің шығындары, у мың.ш.б.	0,9	1,2	1,8	2,2	2,6	2,9	3,3	3,8
Отбасы табысы, х, мың.ш.б.	1,2	3,1	5,3	7,4	9,6	11,8	14,5	18,7

Регрессия параметрлерін табу үшін $\hat{y} = a + b\ln x$, $z = \ln x$ алмастыруын жасаймыз, келесі кестені (1.24-кесте) құрастырамыз ($\varepsilon = y - \hat{y}_x$):

1.24-кесте

Регрессия параметрлерін табу

N	x	z	Y	zy	z^2	y^2	$\hat{y}_x$	ε	ε^2	A_i
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1	1,2	0,182	0,9	0,16	0,03	0,81	0,48	0,42	0,17	46,32
2	3,1	1,131	1,2	1,36	1,28	1,44	1,50	-0,30	0,09	24,91
3	5,3	1,668	1,8	3,00	2,78	3,24	2,07	-0,27	0,07	15,21
4	7,4	2,001	2,2	4,40	4,00	4,84	2,43	-0,23	0,05	10,46
5	9,6	2,262	2,6	5,88	5,12	6,76	2,71	-0,11	0,01	4,22
6	11,8	2,468	2,9	7,16	6,09	8,41	2,93	-0,03	0,00	1,04
7	14,5	2,674	3,3	8,82	7,15	10,89	3,15	0,15	0,02	4,53
8	18,7	2,929	3,8	11,13	8,58	14,44	3,42	0,38	0,14	9,91
Бар-лығы	**71,60**	**15,32**	**18,70**	**41,92**	**35,04**	**50,83**	-	-	**0,57**	**116,6**
Орта-ша	**8,95**	**1,914**	**2,34**	**5,24**	**4,38**	**6,35**	-	-	**0,07**	**14,57**
σ	-	0,85	0,94	-	-	-	-	-	-	-
σ^2	-	0,72	0,89	-	-	-	-	-	-	-

Регрессия теңдеуін табамыз:

$$b = \frac{\overline{zy} - \bar{z}\bar{y}}{\overline{z^2} - (\bar{z})^2} = \frac{5{,}24 - 1{,}914 \cdot 2{,}34}{4{,}38 - (1{,}914)^2} = 1{,}070$$

$$a = \bar{y} - b\bar{z} = 2{,}34 - 1{,}070 \cdot 1{,}914 = 0{,}288,$$

яғни келесі регрессия теңдеуін аламыз: $\hat{y}_x = 0{,}288 + 1{,}070 \cdot \ln x$.

Ары қарай 1.16-кестенің 8–11 бағандарын толтырамыз.

Корреляция индексін келесі формуламен табамыз.

$$\rho_{xy} = \sqrt{1 - \frac{\sigma^2_{\text{калд}}}{\sigma^2_{\text{у}}}} = \sqrt{1 - \frac{0{,}07}{0{,}89}} = 0{,}959,$$

ал детерминация индексі $\rho^2_{xy} = 0{,}92$, ол нәтижелік белгінің вариациясының 92 %-ы фактор белгінің вариациясымен түсіндіріледі, ал 8 %-ы басқа факторлар үлесіне тиеді.

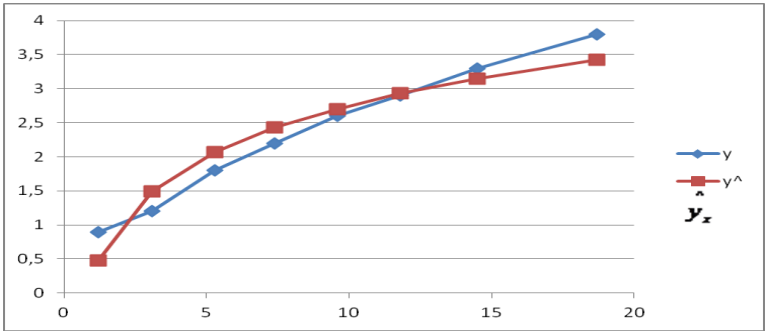
Аппроксимацияның орташа қатесі $\bar{A} = 14{,}57\%$, бұл мүмкін емес жоғары болады.

Фишердің F-критерийі:

$$F = \frac{\rho_{xy}^2}{1-\rho_{xy}^2} \cdot \frac{n-m-1}{m} = \frac{0{,}920}{1-0{,}920} \cdot \frac{8-1-1}{1} = 68{,}16 \cdot$$

Нақты мән кестелік мәннен артық F=5,99.

Графикте кіріс мәліметтерді және регрессия сызығын көрсетеміз (1.9-сурет).



1.9-сурет. Регрессия сызығын құру

1.25-кесте

Шешуі

	x	z	y	Zy	z^2	y^2	Y^	ε	ε^2	Ai	(y-yорт)^2	(z-zорт)^2
1	1,2	0,182	0,9	0,16	0,03	0,81	0,483	0,42	0,17	46,29%	2,07	3,00
2	3,1	1,131	1,2	1,36	1,28	1,44	1,499	-0,30	0,09	24,95%	1,29	0,61
3	5,3	1,668	1,8	3,00	2,78	3,24	2,073	-0,27	0,07	15,19%	0,29	0,06
4	7,4	2,001	2,2	4,40	4,00	4,84	2,431	-0,23	0,05	10,49%	0,02	0,01
5	9,6	2,262	2,6	5,88	5,12	6,76	2,709	-0,11	0,01	4,21%	0,07	0,12
6	11,8	2,468	2,9	7,16	6,09	8,41	2,930	-0,03	0,00	1,04%	0,32	0,31
7	14,5	2,674	3,3	8,82	7,15	10,89	3,151	0,15	0,02	4,52%	0,93	0,58
8	18,7	2,929	3,8	11,13	8,58	14,44	3,423	0,38	0,14	9,92%	2,14	1,03
Сом-масы	**71,600**	**15,315**	**18,700**	**41,918**	**35,035**	**50,830**	**18,70**	**0,000**	**0,568**	**1,166**	**7,119**	**5,717**
Сред-нее	**8,950**	**1,914**	**2,338**	**5,240**	**4,379**	**6,354**	**2,34**	**0,000**	**0,071**	0,146	0,890	0,715
σ	-	0,85	0,94	-	-	-	-	-	-	-	-	-
σ^2	-	0,72	0,89	-	-	-	-	-	-	-	-	-

Бақылау сұрақтары

1.Сызықтық емес теңдеу келтіріңіз.

2. Параметр бойынша сызықтықеместікті қалай жою болады?

3. Функцияның қандай түрлерін білесіз?

4. Өндірістік функция қай кезде қолданылады және неліктен?

5. Регрессиялық модельдердің әртүрлері үшін икемділік коэффициенттері.

6. Сызықтық емес регрессия үшін Фишердің F-критерийі және корреляция.

1.14. Кобба-Дугластың өндірістік функциясы

Макроэкономикалық өндірістік функция – бұл *У* өнім көлемі, К капиталды шығындар және L еңбек шығындары арасындағы статистикалық мәні бар байланыс. Макро және микроэкономикалық деңгейде есепті модельдеу және шешу үшін Кобба-Дугластың өндірістік функциясы (КД) қолданылады:

$$Y = AK^{\alpha}L^{\beta}\varepsilon,$$

мұндағы А, α,β – функцияның параметрлері, А>0; 0<α<1; 0<β<1, ε – кездейсоқ мүше.

1927 жылы Пол Дуглас, экономист *У* нақты өнім көлемінің, К капиталды шығындардың және L еңбек шығындардың көрсеткіштерінің логарифмдерінің графиктерін бір диаграммаға салсақ, онда капитал және еңбек шығындарының көрсеткіштерінің графиктерінің нүктеден нүктеге дейінгі арақашықтығы тұрақты пропорцияны құрайтынын байқады [3,9,16].

Математик Чарльз Кобб бұл ерекшеліктің математикалық тәуелділігін анықтады және $Y = AK^{\alpha}L^{\beta}$ функциясын ұсынды. Мұндай функцияны 30-ға жуық жыл бұрын Филипп Уикстид ұсынған болатын, бірақ Кобб және Дуглас бұл функцияны

тұрғызу үшін эмпирикалық мәліметтерді бірінші болып пайдаланды.

Уақыттық қатарлардың мәліметтерін қолдану арқылы КД функциясын құру кезінде өнімнің шығарылымына техникалық дағдарыстың әсері де ықпал ететінін ескеру керек

Техникалық дағдарыстың әсерін КД функциясын $Y = AK^{\alpha}L^{\beta}e^{rt}\varepsilon$ түрінде жазу арқылы ескеруге болады, мұндағы t – уақыт, r – техникалық дағдарыстың әсеріне байланысты өнімнің өсу қарқыны.

Кобба-Дугластың өндірістік функциясының қасиеттер.

Өнім шығарылымының икемділігі.

Өнім шығарылымының капитал және еңбек бойынша икемділігі α және β: $\frac{\partial Y/Y}{\partial K/K} = \alpha$; $\frac{\partial Y/Y}{\partial L/L} = \beta$.

Бұл капитал шығындарының 1 %-ға арттуы өнім шығарылымының α %-ға өсуіне алып келетінін, ал еңбек шығындарының 1 %-ға арттуы өнім шығарылымының β %-ға өсуіне алып келетінін көрсетеді.

1) Өндірістің масштабына байланысты эффект.

K,L факторларының әрқайсысының шығындарының **λ** есе өсуі кезінде, шығарылым $\lambda^{\alpha+\beta}$ есе өседі.

Бұл былай түсіндіріледі:

– егер α+β>1, онда КД функциясының өндірістің масштабына байланысты өспелі берілімі болады;

– егер α+β<1, онда КД функциясының өндірістің масштабына байланысты кемімелі берілімі болады;

– егер α+β=1, онда КД функциясының өндірістің масштабына байланысты тұрақты берілімі болады.

3) Өндірістік факторлардың болжанатын үлестері.

Нарықтық экономикада α және β бағалары капитал және еңбек есесінен алынған табыстың болжанатын үлестері ретінде интерпретацияланады. Жиынтық сызықтық регрессия моделі көмегімен Кобба-Дугластың өндірістік функциясының параметрлерін бағалау үшін функцияны логарифмдейміз:

$$\ln Y = \ln A + \alpha \ln K + \beta \ln L + \ln \varepsilon .$$

Бұл кезде сызықтық регрессиялық модельді бағалау үшін қажетті *lnε* қателіктерінің қасиеттері бар деп қарастырылған. *Y, K, L* мәліметтерінің қатарлары бойынша олардың логарифмдері есептелінеді және олар үшін регрессия теңдеуі бағаланылады.

1.8-мысал. Y Өндіріс көлемі жайлы мәліметтер берілген (ш.б), олар капитал шығындары және 12 жыл бойғы L мемлекеттің еңбек шығындары (1.26-кесте).

1.26-кесте

Бастапқы мәліметтер

T	Y	K	L	T	Y	K	L
1	100	100	100	7	153	216	145
2	112	114	110	8	184	236	154
3	124	131	123	9	189	266	154
4	143	149	125	10	227	335	196
5	151	176	138	11	218	397	193
6	155	198	140	12	179	417	147

“Регрессия” компьютерлік бағдарламасын қолдана отырып регрессияның бағалық теңдеуі алынды: $\ln \widetilde{Y} = -0{,}302 + 0{,}15 \ln K + 0{,}92 \ln L$ яғни бағалау α=0,15, β=0,92. Бұл капитал шығындарының 1 % өсуі өнім шығарылымының 0,15 % өсуіне, ал еңбек шығындарың 1 % өсуі – өнім шығарылымын 0,92 % өсуіне алып келетіндігін көрсетеді.

Бақылау сұрақтары

1. Кобба-Дугластың өндірістік функциясының қасиеттерін айтыңыз.
2. Кобба-Дугластың өндірістік функциясының қасиеттерін сипаттаңыз.
3. Өнім шығарылымының капитал мен еңбекке қатысты икемділігі қандай?

1.15. Жиынтық регрессия және корреляция

Модельдеу кезінде зерттеу объектісіне әсер ететін факторлардың әсерін ескерсек, жұптық регрессия жақсы нәтиже бере алады. Егер бұл әсерді ескеру керек болса, онда бұл жағдайда басқа факторлардың әсерін оларды модельге енгізе отырып айқындауға тырысу керек, яғни көптік регрессия теңдеуін құру керек [1,3,6,9,12]:

$$y = f(x_1, x_2, ..., x_m) + \varepsilon,$$

мұндағы y – тәуелді айнымалы (нәтитижелік мән), x_i – тәуелсіз, немесе түсіндіретін айнымалылар (мәндер-факторлар). Жиынтық регрессия сұраныстың, акциялардың табысының, өндірістің шығындар функциясын зерттеуінің мәселелерін шешуде, макроэкономикалық есептеулерде және эконометриканың басқа да мәселелерінде кең қолданылады.

Қазіргі кезде жиынтық регрессия – эконометрикада кең тараған әдістердің бірі. Жиынтық регрессияның негізгі мақсаты – факторлардың әрқайсысының әсерін арнайы, және модельденетін көрсеткішке олардың жиынтық әсерін анықтап модель құру.

Модельдің спецификациясы. Жиынтық регрессия теңдеуін құру кезінде факторларды таңдау.

Жиынтық регрессия теңдеуін құру модель спецификациясы туралы сауалды шешуден басталады. Оған сұрақтардың екі шеңбері кіреді: факторларды таңдау және регрессия теңдеуінің түрін таңдау [12].

Жиынтық регрессия теңдеуіне факторлардың әйтеуір бір тізімін енгізу ең алдымен зерттеушінің модельденетін көрсеткіштің басқа экономикалық құбылыстармен өзара байланыстың табиғаты туралы көзқарасына байланысты. Жиынтық регрессияға енгізілетін факторлар келесі талаптарға сай болуы керек:

1. Олар мөлшерлі өлшенетін болуы керек. Егер модельге мөлшерлі өлшемі жоқ сапалы факторды енгізу қажет болса, онда оған мөлшерлі нақтылық беру қажет.

2. Факторлар интеркоррелированланған болмауы керек және олардың нақты функционалдық байланысы болмауы керек.

Модельге жоғары интеркорреляциясы бар факторларды енгізу көзделмеген нәтижелерге әкелуі мүмкін – нормальді теңдеулер жүйесі нашар негізделген болуы мүмкін және мұнда регрессия коэффициенттері бағаларының тұрақсыздығы және сенімсіздігі пайда болуы мүмкін. Егер факторлар арасында жоғары корреляция болса, онда олардың нәтижелік көрсеткішке әсерін анықтай алмаймыз және регрессия теңдеуінің параметрлері интерпретацияланбайтын болып шығады.

Жиынтық регрессия теңдеуіне енгізілетін факторлар тәуелсіз айнымалының вариациясын түсіндіру керек. Егер m факторлар тізімі бар модель құрастырылса, онда ол үшін R^2 детерминация көрсеткіші есептелінеді, ол регрессияда қарастырылатын m факторлар есесінен нәтижелік көрсеткіштің түсіндірілген вариацияның үлесін тіркейді. Модельде ескерілмей қалған факторлардың әсері тиісті S^2 қалдық дисперсиясы бар 1- R^2 ретінде бағаланылады.

Регрессияға қосымша m+1 фактор енгізгенде, детерминация коэффициенті өсу керек, ал қалдық дисперсия азаюы керек:

$$R^2_{m+1} > R^2_m \text{ и } S^2_{m+1} \le S^2_m .$$

Егер осылай болмаса және берілгендер бір-бірінен айырмашылығы болмаса, онда талдауға енгізілген x_{m+1} фактор модельді жақсартпайды және іс-жүзінде артық фактор болып табылады.

Осылайша, теорияда регрессиялық модель факторлардың кез келген санын ескеруге мүмкіндік бергенімен, іс жүзінде мұның қажеттігі жоқ.

Факторларды таңдау сапалы теориялық-экономикалық талдау негізінде жасалады. Бірақ теориялық талдау қарастырылатын мәндердің мөлшерлі өзарабайланысы туралы және факторды модельге енгізу дұрыстығы туралы сұраққа дәл жауап бере алмайды. Сондықтан факторларды таңдау екі кезеңде жүзеге асырылады:

Бірінші кезеңде факторлар мәселенің мәніне қарай таңдалады, екінші кезеңде – регрессияның параметрлері үшін корреляцияның көрсеткіштерінің матрицасы негізінде статистика анықталады.

Интеркорреляция коэффициенттері (яғни түсіндіретін айнымалылар арасындағы корреляциялар) модельден егіздейтін факторларды алып тастауға мүмкіндік береді. Екі айнымалы коллинеарлы, яғни өзара сызықты тәуелділікте болады деп есептеледі, егер $r_{x_i x_j} \geq 0{,}7$. Егер факторлар коллинеарлы болса, онда олар бірін бірі егіздейді және олардың біреуін регрессиядан алып тастау ұсынылады.

Мұнда таңдау нәтижемен тығыз байланысты факторға емес, нәтижемен жеткілікті тығыз байланыс кезінде басқа факторлармен байланыстың ең аз тығыздығы бар факторға жасалынады. Бұл талапта жиынтық регрессияның арналымы олардың бір-бірінен тәуелсіздік жағдайында факторлардың кешенді әсер ету әдісі ретінде көрінеді.

Жиынтық регрессияның аппаратын қолдану кезінде факторлардың мультиколлинеарлылығы кезінде үлкен қиындықтар кездеседі, бұл кезде екіден артық факторлар өзара сызықтық тәуелділікпен байланысты, яғни мұнда факторлардың бір-біріне жиынтық әсері орын алады. Факторлардың мультиколлинеарлылығының болуы кейбір факторлар әрқашан унисонда қимылдайтынын білдіреді. Нәтижесінде вариация шығыс мәліметтерде толығымен тәуелсіз болуын тоқтатады және мұнда әрбір фактордың әсерін арнайы бағалай алмаймыз.

Модельге мультиколлинеарлық факторларды енгізу келесі нәтижелер болатындықтан керексіз:

1. Жиынтық регрессияның параметрлерінің интерпретациясы факторлардың «таза» түріндегі әсерінің сипаттамасы ретінде қиындайды, өйткені факторлар коррелированған; сызықтық регрессияның параметрлерінің экономикалық мәні жойылады.

2. Параметрлердің бағалары сенімді емес, үлкен стандартты қателіктер анықталынады және зерттеулер көлемінің өзгерісімен өзгереді (мөлшері бойынша ғана емес, таңбасы бойынша да), бұл модельді талдау және болжау үшін қажетсіз етеді.

Факторлардың мультиколлинеарлылығын бағалау үшін факторлар арасындағы корреляцияның жұп коэффициенттерінің матрицасының анықтаушысы қолданылады.

Егер факторлар өзара коррелирленбеген болса, онда факторлар арасындағы корреляцияның жұп коэффициенттерінің матрицасы бірлік матрица болар еді, өйткені барлық диагональді емес элементтер $r_{x_i x_j}$ ($i \neq j$) нөлге тең болар еді. Осылайша, үш түсіндіруші айнымалысы бар теңдеу үшін

$$\hat{y} = a + b_1 x_1 + b_2 x_2 + b_3 x_3$$

факторлар арасындағы корреляцияның жұп коэффициенттерінің матрицасының бірге тең анықтаушысы болар еді:

$$\det R = \begin{vmatrix} r_{x_1x_1} & r_{x_1x_2} & r_{x_1x_3} \\ r_{x_2x_1} & r_{x_2x_2} & r_{x_2x_3} \\ r_{x_3x_1} & r_{x_3x_2} & r_{x_3x_3} \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1$$

Егер, керісінше, факторлар арасында толық сызықтық тәуелділік болса және корреляцияның барлық коэффициенттері бірге тең болса, онда бұл матрицаның анықтаушысы нөлге тең:

$$\det R = \begin{vmatrix} r_{x_1x_1} & r_{x_1x_2} & r_{x_1x_3} \\ r_{x_2x_1} & r_{x_2x_2} & r_{x_2x_3} \\ r_{x_3x_1} & r_{x_3x_2} & r_{x_3x_3} \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 0$$

Фактораралық корреляцияның матрицасының анықтаушысы нөлге жақын болған факторлардың мультиколлинеарлығы күштірек және жиынтық регрессияның нәтижелері сенімдірек болады. Және, керісінше, фактораралық корреляцияның матрицасының анықтаушысы бірге жақын болған факторлардың мультиколлинеарлығы азырақ болады.

Күшті фактораралық корреляцияны жеңудің бірқатар тәсілдері бар. Мультиколлинеарлықты жоюдың ең қарапайым түрі модельден бір немесе бірнеше факторларды жою.

Факторлардың ішкі корреляциясын есептеу жолдарының бірі регрессияның біріккен теңдеулеріне ауысу, яғни

факторлардың тек әсерін ғана емес, олардың өзараәрекетін де көрсететін теңдеулер.

Осылайша, егер $y=f(x1,x2, x3)$, онда келесі біріккен теңдеуді құруға болады:

$$y = a + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + \varepsilon.$$

Қарастырылатын теңдеу бірінші реттегі (екі фактордың өзара әрекеті) өзара әрекетінен тұрады. Модельге бұдан жоғары реттегі өзара әрекеттерді енгізуге болады, егер Фишердің *F*-критерийі бойынша олардың статистикалық мәні дәлелденсе. Бірақ, негізінен, үшінші және бұдан жоғары реттегі өзара әрекеттер статистикалы мәнсіз болып саналады.

Регрессияға енгізілетін факторларды таңдау регрессияның әдістерін іс жүзінде қолданудың маңызды кезеңдерінің бірі болып табылады.

Факторларды корреляция көрсеткіштері негізінде таңдау тәсілдемелері әртүрлі болады. Олар жиынтық регрессия теңдеулерін құруды әртүрлі әдістемелерге алып келеді. Регрессия теңдеуін құруға қандай әдістеме қабылданғанына байланысты оны ЭЕМ-да шешу алгоритмі де өзгереді.

Жиынтық регрессия теңдеуін құрудың келесі әдістері кең қолданылады:

1. Алып тастау әдісі – факторларды олардың толық тізімінен қарау.

2. Енгізу әдісі – факторды қосымша енгізу.

3. Қадамдық регрессиялық талдау – алдында енгізілген факторды алып тастау.

Факторларды таңдау кезінде келесі ережені қолдану ұсынылады: көбінесе енгізілетін факторлардың саны регрессия құрастырылатын жиынның көлемінен 6–7 есе аз болады. Егер бұл қатынас бұзылса, онда қалдық дисперсиясының еркіндік дәрежесінің саны өте аз болады. Бұл регрессия теңдеуінің параметрлерінің статистикалық мәні жоқтығына алып келеді, ал бұл кезде *F*-критерий кестелік мәннен аз болады.

Бақылау сұрақтары

1. Сызықтық жиынтық регрессия теңдеуін келтіріңіз.
2. Жиынтық регрессия теңдеуін құру кезінде факторларды таңдау.
3. Көптік регрессияның қасиеттері жұптық регрессияның қасиеттерінен қалай ерекшеленеді?
4. Мультиколлинеарлық құбылысына анықтама беріңдер.
5. Мультиколлинеарлықты жою үшін қандай шараларды қолдануға болады.

1.15.1. Ең кіші квадраттар әдісі (ЕКӘ). ЕКӘ негізінде алынатын бағалардың қасиеттері

Жиынтық регрессия теңдеуінің келесі түрлері болуы мүмкін: сызықтық және сызықтық емес.

Параметрлердің нақты интерпретациясы есесінен сызықтық функция көп қолданылады. Сызықтық жиынтық регрессияда

$$\hat{y}_x = a + b_1 x_1 + b_2 x_2 + ... + b_m x_m$$

x кезіндегі параметрлер «таза» регрессияның коэффициенттері деп аталады. Ол орташа деңгейде бекітілген басқа факторлардың өзгертілмеген мәні кезінде тиісті фактордың бір бірлікке өзгерген кезде нәтиженің орташа өзгерісін сипаттайды.

Сызықтық жиынтық регрессия моделін қарастырайық

$$y = a + b_1 x_1 + b_2 x_2 + ... + b_m x_m + \varepsilon. \qquad (1.52)$$

Сызықтық жиынтық регрессия моделінің параметрлерін бағалауының классикалық тәсілдемесі ең кіші квадраттар әдісіне негізделген. ЕКӘ y нәтижелік мәннің нақты мәндерінің $\hat{y}$ есептелінетін мәндерінен ауытқуларының квадраттарының

сомасы минималды болатын параметрлердің бағаларын алуға мүмкіндік береді:

$$\sum_i (y_i - \hat{y}_{x_i})^2 \to \min .$$

Математикалық талдау курсынан білетініміз, бірнеше айнымалылардан тұратын функцияның экстремумын табу үшін параметрлердің әрқайсысы бойынша бірінші реттегі туындыларын тауып, оларды нөлге теңестіруіміз керек.

Сонымен, бізде m +1 аргументі бар функциямыз бар:

$$S(a, b_1, b_2, ..., b_m) = \sum (y - a - b_1 x_1 - b_2 x_2 - ... - b_m x_m)^2$$

бірінші реттегі жеке туындыларын табамыз:

$$\frac{\partial S}{\partial a} = -2\sum (y - a - b_1 x_1 - b_2 x_2 - ... - b_m x_m) = 0;$$

$$\frac{\partial S}{\partial b_1} = -2\sum x_1 (y - a - b_1 x_1 - b_2 x_2 - ... - b_m x_m) = 0;$$

.....

$$\frac{\partial S}{\partial b_m} = -2\sum x_m (y - a - b_1 x_1 - b_2 x_2 - ... - b_m x_m) = 0.$$

Қарапайым түрлендірулерден кейін жиынтық регрессияның сызықтық теңдеуінің (1.52) параметрлерін табу үшін сызықтық нормальді теңдеулер жүйесін аламыз.

$$\begin{cases} na + b_1 \sum x_1 + b_2 \sum x_2 + ... + b \sum x_m = \sum y \\ a\sum x_1 + b_1 \sum x_1^2 + b_2 \sum x_1 x_2 + ... + b_m \sum x_1 x_m = \sum y x_1, \\ ..... \\ a\sum x_m + b_1 \sum x_1 x_m + b_2 \sum x_2 x_m + ... + b_m \sum x_m^2 = \sum y x_m \end{cases}$$

Екіфакторлы модель үшін берілген жүйе келесі түрде болады:

$$\begin{cases} na + b_1 \sum x_1 + b_2 \sum x_2 = \sum y, \\ a \sum x_1 + b_1 \sum x_1^2 + b_2 \sum x_1 x_2 = \sum y x_1, \\ a \sum x_2 + b_1 \sum x_1 x_2 + b_2 \sum x_2^2 = \sum y x_2. \end{cases}$$

Ең кіші квадраттар әдісі стандарталған масштабтағы жиынтық регрессия теңдеуіне де қолданылады:

мұнда

$$t_y = \beta_1 t_{x_1} + \beta_2 t_{x_2} + ... + \beta_m t_{x_m} + \varepsilon$$

$t_y, t_{x_1}, ..., t_{x_m}$ – стандартталған айнымалылар: $t_y = \dfrac{y - \overline{y}}{\sigma_y}$, $t_{x_i} = \dfrac{x_i - \overline{x_i}}{\sigma_{x_i}}$, олардың орташа мәні нөлге тең: $\overline{t_y} = \overline{t_{x_i}} = 0$, ал орташа квадраттық ауытқуы бірге тең: $\sigma_{t_y} = \sigma_{t_{x_i}} = 1; \beta_i$ – регрессияның стандартталған коэффициенттері [12].

Регрессияның стандартталған коэффициенттері тиісті x_i факторы басқа факторлардың өзгермейтін орташа деңгейде бір бірлікке өзгерсе, орта есеппен нәтиже қанша бірлікке өзгеретінін көрсетеді.

Айнымалылар орталықтандырылған және нормаланған болып берілгендіктен, β_i регрессияның стандартталған коэффициенттерін өзара салыстыруға болады. Оларды бір-бірімен салыстыра отырып, факторларды нәтижеге әсер ету күшіне байланысты ранжирлеуге болады. Бұл регрессияның стандартталған коэффициенттерінің «таза» регрессия коэффициенттерінен айыратын артықшылығы.

Ең кіші квадраттар әдісін стандартталған масштабтағы жиынтық регрессия теңдеуіне қолданып, келесі теңдеулер жүйесін аламыз:

$$r_{yx_1} = \begin{cases} r_{yx_1} = \beta_1 + \beta_2 r_{x_1x_2} + \beta_3 r_{x_1x_3} + ... + \beta_m r_{x_1x_m}, \\ r_{yx_2} = \beta_1 r_{x_1x_2} + \beta_2 + \beta_3 r_{x_1x_3} + ... + \beta_m r_{x_1x_m}, \\ ..... \\ r_{y_m} = \beta_1 r_{x_1x_m} + \beta_2 r_{x_2x_m} + \beta_3 r_{x_3x_m} + ... + \beta_m, \end{cases}$$

мұндағы r_{yx_m} және $r_{x_ix_j}$ – жұптық және фактораралық корреляция коэффициенттері.

b_i «таза» регрессия коэффициенттері β_i регрессияның стандартталған коэффициенттерімен байланысы келесіде:

$$b_i = \beta_i \frac{\sigma_y}{\sigma_{x_i}}.$$

Сондықтан стандартталған масштабтағы регрессия теңдеуінен (1.55) айнымалылардың натуралды масштабтағы регрессия теңдеуіне (1.52) ауысуға болады, бұл кезде a параметрі келесідей анықталады: $a = \bar{y} - b_1\bar{x}_1 - b_2\bar{x}_2 - ... - b_m\bar{x_m}$.

Регрессияның стандартталған коэффициенттерінің қарастырылған мағынасы факторларды таңдау кезінде қолдануға мүмкіндік береді –модельден ең кіші β_i мәні бар факторлар жойылады.

Жиынтық регрессияның сызықтық теңдеуі негізінде

$$y = a + b_1x_1 + b_2x_2 + ... + b_mx_m + \varepsilon$$

Регрессияның жеке теңдеулері анықталуы мүмкін:

$$\begin{cases} \hat{y}_{x_1,x_2,x_3,...x_m} = f(x_1), \\ \hat{y}_{x_2x_1,x_3...,x_m} = f(x_2), \\ ..... \\ \hat{y}_{x_mx_1x_x,...x_{m-1}} = f(x_m). \end{cases}$$

Яғни орташа деңгейде қалған факторларды бекіту кезінде тиісті x_i факторы бар нәтижелік мәнмен байланысты регрессия теңдеулері. Осыны ашып жазсақ, (1.59) жүйені келесі түрде жазуға болады:

$$\begin{cases} y_{x_1x_2x_3,\ldots x_m} = a + b_1x_1 + b_2\bar{x}_2 + b_3\bar{x}_3 + \ldots + b_m\overline{x_m} + \varepsilon, \\ y_{x_2x_1x_3,\ldots x_m} = a + b_1\bar{x}_1 + b_2x_2 + b_3\bar{x}_3 + \ldots + b_m\bar{x}_m + \varepsilon, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_{x_mx_1x_2,\ldots x_{m-1}} = a + b_1\bar{x}_1 + b_2\overline{x_2} + b_3\bar{x}_3 + \ldots + b_mx_m + \varepsilon. \end{cases}.$$

Бұл теңдеулерге тиісті факторлардың орташа мәндерін қойсақ, олар сызықтық регрессияның жұптық теңдеулерінің түріне ие болады, яғни

$$\begin{cases} \hat{y}_{x_1,x_2,x_3,\ldots x_m} = A_1 + b_1x_1, \\ \hat{y}_{x_2x_1,x_3\ldots x_m} = A_2 + b_2x_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \hat{y}_{x_mx_1x_x,\ldots x_{m-1}} = A_m + b_mx_m. \end{cases}.$$

Мұнда

$$\begin{cases} A_1 = a + b_2\overline{x_2} + b_3\overline{x_3} + \ldots + b_m\bar{x}_m, \\ A_2 = a + b_1\overline{x_1} + b_3\bar{x}_3 + \ldots + b_m\overline{x_m}. \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ A_m = a + b_1\overline{x_1} + b_2\overline{x_2} + b_3\overline{x_3} + \ldots + b_{m-1}\bar{x}_{m-1}. \end{cases}$$

Жұптық регрессияға қарағанда регрессияның жеке теңдеулері фактордың нәтижеге жеке әсерін сипаттайды, өйткені басқа факторлар өзгермейтін деңгейде бекітілген. Басқа факторлардың әсер ету тиімділігі жиынтық регрессияның теңдеуіндегі бос мүшеге біріктірілген. Бұл регрессияның жеке

теңдеулері негізінде икемділіктің жеке коэффициенттерін анықтауға мүмкіндік береді:

$$Э_{y_{x_i}} = b_i \frac{x_i}{\hat{y}_{x_i x_1, x_2, \ldots x_{i-1}, x_{i+1}, \ldots x_m}},$$

мұндағы b_i – жиынтық регрессияның теңдеуіндегі x_i факторы үшін регрессия коэффициенті, $\hat{y}_{x_i x_1 x_2 \ldots x_{i-1}, x_{t+1}, \ldots x_m}$ – регрессияның жеке теңдеуі.

Икемділіктің жеке коэффициенттерімен қатар жиынтық бойынша орташа икемділік көрсеткіштері анықталуы мүмкін:

$$\overline{Э_i} = b_i \frac{\overline{x_i}}{\overline{y}_{x_i}}.$$

Олар тиісті фактор 1 %-ға өзгергенде нәтиже орта есеппен қанша пайызға өзгеретінін көрсетеді. Икемділіктің орташа көрсеткіштерін бір-бірімен салыстыруға болады және оларды нәтижеге әсер ету күшіне байланысты ранжирлеуге болады.

1.25-тапсырма. Сату кәсіпорын жүйесі бар болады, 12 дүкеннің тұрақты мәліметтері және оның жұмысы 1.27-кестеде көрсетілген [8,9,10].

1.27-кесте

Дүкендер №	Тауар айналымы, млн. теңге	Сату аймағы, мың м^2	Орташа нәтижелі көрсеткіштер
1	19,76	0,24	8,25
2	38,09	0,31	10,24
3	40,95	0,55	9,31
4	41,08	0,48	11,01
5	56,29	0,78	8,54
6	68,51	0,98	7,51
7	75,01	0,94	12,36
8	89,05	1,21	10,81
9	91,13	1,29	9,89

10	91,26	1,12	13,72
11	99,84	1,29	12,27
12	108,55	1,49	13,92

Диаграмманы тұрғызу керек және жылдық тауар айналымы (у) сату аймағына тәуелді болады (x_1) және (x_2) орташа нәтижелі көрсеткіштер арасындағы байланыс формуласын анықтау және (x_1) және (x_2) әр факторын.

1.26-тапсырма. 1.27-кестеде берілген ақпарат негізінде (x_1) магазиннің сату алаңына тәуелді 1 жылдық тауар айналымының 2 факторлық құрылған және (x_2), бір күн ішіндегі сатып алушылар орта саны келесі түрдегі

$$\hat{y}_t = -10{,}8153 + 61{,}6583x_{1t} + 2{,}2748x_{2t}.$$

Анықтау керек:

1. Теңделген регрессия коэффициентіне экономикалық түсіндіру беру.

2. 1.27-кестенің мәліметтері негізінде қатысушылардың саудалық ауданнан және орташа санынан жылдық тауар айналымының икемділіктің эмпирикалық коэффициенттерін есептеу.

3. Регрессия теңдеулерінің негізінде қатысушылардың саудалық ауданнан және орташа санынан жылдық тауар айналымының икемділіктің жеке коэффициенттерін бағалау.

Бақылау сұрақтары

1. Жиынтық регрессия теңдеуінің параметрлерін бағалау.
2. Жиынтық корреляция.
3. Қандай жағдайда регрессия коэффициенттері дұрыс болады?
4. Түсіндіруші айнымалылар арасындағы байланыс қалай анықталады?

1.15.2. Факторлардың мәнділігін тексеру және регрессия сапасының көрсеткіштері

Жиынтық регрессия теңдеуінің іс жүзіндегі мәнділігі жиынтық корреляция және оның квадраты – детерминация көрсеткіші көмегімен бағаланады.

Жиынтық корреляция көрсеткіші зерттелетін қасиеті бар факторлардың қарастырылатын тізімінің байланыс тығыздығын сипаттайды, немесе факторлардың нәтижеге біріккен әсер ету тығыздығын бағалайды [3,12,15].

Жиынтық корреляция көрсеткіші байланыс формасына тәуелсіз жиынтық корреляция индексі ретінде табылады:

$$\mathrm{R}_{\mathrm{y}x_1x_2\ldots x_m} = \sqrt{1-\frac{\sigma^2_{\text{калд}}}{\sigma^2_{\mathrm{y}}}},$$

мұндағы σ_y^2 – нәтижелік мәннің жалпы дисперсиясы; $\sigma_{кал\partial}^2$ – қалдық дисперсия.

Жиынтық корреляция индексінің өзгеру шегі 0-ден 1-ге дейін. Оның мәні 1-ге жақын болған сайын, нәтижелік мәннің зерттелетін факторлар арасындағы байланыс тығызырақ. Жиынтық корреляция индексінің мөлшері корреляцияның максималды жұптық индексінен артық немесе тең болуы керек:

$$\mathrm{R}_{\mathrm{y}x_1x_2\ldots x_m} \geq r_{yx_i\,(\max)} \quad (\mathrm{i}=\overline{1,m}).$$

Регрессиялық модельге факторларды дұрыс енгізгенде жиынтық корреляция индексі жұптық тәуелділіктің корреляция индексінен өзгешілігі үлкен болады. Егер жиынтық регрессия теңдеуіне қосымша енгізілген факторлар үшінші реттегі болса, онда жиынтық корреляция индексі іс-жүзінде жұптық корреляция индексіне сәйкес келуі мүмкін. Осы жерден жиынтық және жұптық корреляция индекстерін салыстырсақ, регрессия теңдеуіне әйтеуір – бір факторды енгізу дұрыстығы түсінікті болады.

Жиынтық корреляция индексін есептеу жиынтық регрессия теңдеуін және оның негізінде қалдық дисперсияны анықтауды көздейді:

$$\sigma^2_{\text{калд}} = \frac{1}{n}\sum (y - \hat{y}_{x_1 x_2 \ldots x_m})^2.$$

Жиынтық детерминация индексінің келесі формуласын қолдануға болады:

$$\mathrm{R}^2_{yx_1x_2\ldots x_m} = 1 - \frac{\sum (y - \hat{y}_{x_1 x_2 \ldots x_m})}{\sum (y - \bar{y})^2}.$$

Мәндердің сызықтық тәуелділігі кезінде Жиынтық корреляция индексінің формуласын келесі түрде беруге болады:

$$\mathrm{R}_{yx_1x_2\ldots x_m} = \sqrt{\sum \beta_i r_{yx_i}},$$

мұндағы β_i – регрессияның стандартталған коэффициенттері;

r_{yx_i} – әрбір фактор сайын нәтиженің корреляциясының жұптық коэффициенттері.

Сызықтық регрессия үшін жиынтық корреляция индексінің формуласы *жиынтық корреляцияның сызықтық коэффициенті* немесе *корреляцияның жиынтық коэффициенті* деп аталады.

Корреляцияның жиынтық коэффициентін сызықтық тәуелділік кезінде корреляцияның жұптық коэффициенттерінің матрицасы:

$$\mathrm{R}_{yx_1x_2\ldots x_m} = \sqrt{1 - \frac{\Delta r}{\Delta r_{11}}},$$

мұнда

$$\Delta r = \begin{vmatrix} 1 & r_{yx_1} & r_{yx_2} & \ldots & r_{yx_p} \\ r_{yx_1} & 1 & r_{x_1x_2} & \ldots & r_{x_1x_p} \\ r_{yx_2} & r_{x_2x_1} & 1 & \ldots & r_{x_2x_p} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ r_{yx_p} & r_{x_px_1} & r_{x_px_2} & \ldots & 1 \end{vmatrix}$$

– корреляцияның жұптық коэффициенттерінің матрицасының анықтаушысы;

$$\Delta r_{11} = \begin{vmatrix} 1 & r_{x_1x_2} & \dots & r_{x_1x_p} \\ r_{x_2x_1} & 1 & \dots & r_{x_2x_p} \\ \dots & \dots & \dots & \dots \\ r_{x_px_1} & r_{x_px_2} & \dots & 1 \end{vmatrix}$$

– фактораралық корреляцияның матрицасының анықтаушысы.

Корреляция жиынтық коэффициентінің мөлшері факторлардың әрқайсысымен жасалынған нәтиженің корреляциясына ғана емес, фактораралық корреляцияға байланысты. Қарастырылған формула корреляцияның жиынтық коэффициентін есептейді.

Жиынтық корреляцияның қарастырылған көрсеткіштерінде (индекс және коэффициент) қалдық дисперсия қолданылады, оның азаюы жағына қарай жүйелік қателігі бар, зерттеулердің берілген *n* көлемі кезінде регрессия теңдеуінде анықталады, параметрлер көп болған сайын оның мәні көбейеді. Егер x_i кезінде параметрлер саны *m* болса және зерттеулердің санына жақындаса, онда қалдық дисперсия нөлге жақын болады және корреляция коэффициенті (индексі) факторлардың нәтижесімен әлсіз байланыс кезінде де бірге жақындайды.

Байланыс тығыздығының көбейіп кетуіне жол бермеу үшін жиынтық корреляцияның реттелген индексі (коэффициенті) қолданылады.

Жиынтық корреляцияның реттелген индексінің еркіндік дәрежесінің санына жасалған түзетулері бар, дәлірек айтқанда, $\sum(y - \hat{y}_{x_1x_2\dots x_m})^2$ квадраттардың қалдық сомасы қалдық вариацияның (*n- m-1)* еркіндік дәрежесінің санына бөлінеді, ал $\sum(y - \bar{y})^2$ ауытқулар квадраттарының қалдық сомасы жалпы *(n-1)* жиыны бойынша еркіндік дәрежесінің санына бөлінеді.

Жиынтық детерминацияның реттелген индексінің формуласы келесі түрде болады:

$$\hat{R}^2 = 1 - \frac{\sum (y - \hat{y})^2 / (n - m - 1)}{\sum (y - \bar{y}) / (n - 1)},$$

мұндағы m – x айнымалысы кезіндегі параметрлер саны; n – зерттеулер саны.

Мұнда

$$\frac{\sum (y - \hat{y}_{x_1 x_2, \ldots x_m})^2}{\sum (y - \bar{y})^2} = 1 - R^2$$

болғандықтан, детерминацияның реттелген индексінің мөлшерін келесі түрде беруге болады:

$$\hat{R}^2 = 1 - (1 - R^2) \cdot \frac{n - 1}{n - m - 1},$$

m мөлшері көп болған сайын, $\hat{R}^2$ және R^2 айырмашылығы көп болады.

Жоғарыда көрсетілгендей, Жиынтық сызықтық регрессияға қатысатын факторларды ранжирлеу регрессияның стандартталған коэффициенттері (β-коэффициенттері) арқылы өткізіле алады. Бұл мақсатқа корреляцияның жеке коэффициенттері көмегімен (сызықтық байланыстар үшін) жетуге болады. Бұдан басқа, корреляцияның жеке көрсеткіштері факторларды таңдау мәселелерін шешуде кең қолданылады: қандай да бір факторды модельге енгізу дұрыстығы корреляцияның жеке көрсеткіштерінің мөлшерімен дәлелдеуге болады.

Корреляцияның жеке коэффициенттері регрессии теңдеуіне енгізілген басқа факторлардың элиминированиесі (әсерін жою) кезінде нәтиже мен тиісті фактор арасындағы байланыс тығыздығын сипаттайды.

Жалпы түрде m факторы бар $y = a + b_1 x_1 + b_2 x_2 \ldots + b_m x_m + \varepsilon$ теңдеуі үшін *у-ке* x_i факторының әсерін өлшейтін корреляцияның жеке коэффициентін келесі формуламен анықталады:

$$r_{yx_1x_1x_2\ldots x_{i-1}x_{i+1}\ldots x_m}=\sqrt{1-\frac{1-R^2_{yx_1x_2\ldots x_i\ldots x_m}}{1-R^2_{yx_1x_2\ldots x_{i-1}x_{i+1}\ldots x_m}}},$$

мұндағы $R^2_{yx_1x_2\ldots x_m}$ – нәтижесі бар барлық m факторлардың детерминациясының Жиынтық коэффициенті; $R^2_{yx_1x_2\ldots x_{i-1}x_{i+1}\ldots x_m}$ – модельге x_i факторын енгізілмеген кездегі детерминация көрсеткіші.

Екі фактор кезінде (1.69,а) формуланың түрі келесідей болады:

$$r_{yx_1x_2}=\sqrt{1-\frac{1-R^2_{yx_1x_2}}{1-r^2_{yx_2}}},\ r_{yx_2x_1}=\sqrt{1-\frac{1-R^2_{yx_1x_2}}{1-r^2_{yx_1}}}.$$

Корреляцияның жеке коэффициентінің реті әсері жойылатын факторлар мөлшерімен анықталады. Мысалы, $r_{yx_1x_2}$ – бірінші реттегі корреляцияның жеке коэффициенті. Осыған сәйкес, жұптық корреляцияның коэффициенттерін нөлдік реттегі коэффициенттер деп аталады. Жеке корреляцияның коэффициенттерін төменірек реттегі жеке корреляцияның коэффициенттері арқылы рекуррентті формула бойынша анықтауға болады:

$$r_{yx_i\cdot x_1x_2\cdot x_{i-1}x_{i+1}\ldots x_m}=\frac{r_{yx_i\cdot x_1x_2\ldots x_{i-1}x_{i+1}x_{m-1}}-r_{yx_m x_1x_2\ldots x_{m-1}}\cdot r_{x_1x_m x_1x_2\ldots x_{i-1}x_{i+1}\ldots x_{m-1}}}{\sqrt{(1-r^2_{yx_m\cdot x_1x_2\ldots x_{m-1}})\cdot(1-r^2_{x_ix_m\cdot x_1x_2\ldots x_{i-1}x_{i+1}\ldots x_{m-1}}}}.$$

Екі фактор кезінде берілген формула келесі түрге келеді:

$$r_{yx_1x_2}=\frac{r_{yx_1}-r_{yx_2}r_{x_1x_2}}{\sqrt{(1-r^2_{yx_2})\cdot(1-r^2_{x_1x_2})}};\ r_{yx_2x_1}=\frac{r_{yx_2}-r_{yx_1}r_{x_1x_2}}{\sqrt{(1-r^2_{yx_1})\cdot(1-r^2_{x_1x_2})}}.$$

Үш факторы бар регрессия теңдеуі үшін екінші реттегі корреляцияның жеке коэффициенттері бірінші реттегі корреляцияның жеке коэффициенттері негізінде анықталады. Осылайша,

$y = a + b_1 x_1 + b_2 x_2 + b_3 x_3 + \varepsilon$ теңдеуі бойынша *i*-нің үш екінші реттегі корреляцияның жеке коэффициенттерін есептеуге болады: $r_{yx_1x_2x_3}, r_{yx_2x_1x_3}, r_{yx_3x_1x_2}$ олардың әрқайсысы рекуррентті формуламен есептеледі. Мысалы, i=1 кезінде $r_{yx_1x_2x_3}$ есептеу формуласы бар:

$$r_{yx_1x_2x_3} = \frac{r_{yx_1x_2} - r_{yx_3x_2} \cdot r_{x_1x_3x_2}}{\sqrt{(1 - r^2_{yx_3x_2})(1 - r^2_{x_1x_3x_2})}}.$$

Рекуррентті формуламен есептелген корреляцияның жеке коэффициенттері -1-ден +1-ге дейін өзгереді, ал формулалар бойынша детерминацияның Жиынтық коэффициенттері -0-ден 1-ге дейін. Оларды бір-бірімен салыстыру факторларды нәтижемен байланысының тығыздығы бойынша ранжирлеуге мүмкіндік береді. Корреляцияның жеке коэффициенттері әрбір фактордың нәтижемен байланыс тығыздығының өлшемін таза түрде береді. Егер

$$t_y = \beta_1 t_{x_1} + \beta_2 t_{x_2} + \beta_3 t_{x_3} + \varepsilon$$

регрессияның стандартталған теңдеуінен $\beta_1 > \beta_2 > \beta_3$ екені көрініп тұрса, яғни нәтижеге әсер ету күші бойынша факторлар реті келесідей: x_1, x_2, x_3, онда факторлардың осы реті корреляцияның жеке коэффициенттерінің қатынасы бойынша анықталады.

Эконометрикада корреляцияның жеке коэффициенттерінің өзіндік мәні жоқ. Оларды модельді құру кезеңінде қолданылады. Осылайша, көпфакторлы модель құрған кезде бірінші кезекте факторлардың толық тізімі бар регрессия теңдеуі анықталады және корреляцияның жеке коэффициенттерінің матрицасы есептеледі. Екінші кезекте жеке корреляция көрсеткішінің мөлшерлі Стьюденттің *t*-критерийі бойынша ең аз және мағынасы жоқ фактор таңдалады. Оны модельден алып тастап, регрессияның жаңа теңдеуі құрылады. Осы процедура барлық корреляцияның жеке коэффициенттері

нөлден өзгеше болғанға дейін жалғасады. Егер мағынасыз фактор алынып тасталса, детерминацияның Жиынтық коэффициенттері регрессиялық модельді құрудың екі көршілес қадамында бір-бірінен айырмашылығы болмайды $R^2_{m+1} \approx R^2_m$, мұнда m – факторлар саны.

Жоғарыда келтірілген формулалардан корреляцияның жеке коэффициенттерінің корреляцияның жиынтық коэффициенттерінің көрсеткіштерімен байланысы көрініп тұр. Корреляцияның жеке коэффициенттерін біле отырып (бірінен соң бірі бірінші, екінші, үшінші және одан жоғары реттегі) корреляцияның жиынтық коэффициентін келесі формула бойынша анықтауға болады:

$$R_{yx_1x_2\ldots x_m} = \sqrt{1-(1-r^2_{yx_1})\cdot(1-r^2_{yx_2x_1})\cdot(1-r^2_{yx_3x_1x_2})\cdot\ldots\cdot(1-r^2_{yx_m\cdot x_1x_2\ldots x_{m-1}}}.$$

Негізінен, екі факторлы теңдеу үшін (1.71) формуланың түрі келесідей болады:

$$R_{yx_1x_2\ldots x_m} = \sqrt{1-(1-r^2_{yx_1})\cdot(1-r^2_{yx_2x_1})}.$$

Нәтижелік мәннің зертелетін фактордан толық тәуелділік кезінде олардың жиынтық әсер ету коэффициенті бірге тең. Бірден талдауға бірінен соң бірі енгізілген факторлармен негізделген $(1-r^2)$ нәтижелік мәннің қалдық вариациясының үлесі шегеріледі. Нәтижесінде түбір астындағы өрнек барлық зерттелетін фактордың жиынтық іс-әрекетін сипаттайды.

Жиынтық регрессия теңдеуінің мәнділігі жұптық регрессиядағыдай Фишердің F-критерийі көмегімен бағаланылады:

$$F = \frac{S_{\text{факт}}}{S_{\text{калд}}} = \frac{R^2}{1-R^2}\frac{n-m-1}{m},$$

мұндағы $S_{факт}$ – еркіндіктің бір дәрежесіне шаққандағы квадраттардың факторлы сомасы; $S_{калд}$ – еркіндіктің бір дәрежесіне шаққандағы квадраттардың қалдық сомасы; R^2 –

Жиынтық детерминация коэффициенті (индексі); m – x айнымалысы кезіндегі параметрлер саны; n – зерттеулер саны.

Мұнда теңдеудің ғана емес, регрессиялық модельге қосымша енгізілген фактор мәнділігі бағаланылады.

Мұндай бағалаудың қажеттігі модельге кірген әрбір фактор нәтижелік мәннің түсіндірілген вариациясын көбейте алмайды. Бұдан басқа, модельде бірнеше фактор болса олар модельге әртүрлі кезеңділікте енгізіле алады. Модельге факторды енгізуді бағалаудың өлшемі жеке *F*-критерий, яғни F_{x_i} болып табылады.

Жеке *F*-критерий регрессиялық модель бойынша еркіндіктің бір дәрежесіне шаққандағы қалдық дисперсиясы бар қосымша енгізілген фактордың әсерімен негізделген факторлы өсімді салыстырумен тұрғызылды. Жалпы түрде x_i факторы үшін жеке *F*-критерий келесідей анықталады

$$F_{x_i} = \frac{R^2_{yx_1 \ldots x_i \ldots x_m} - R^2_{yx_1 \ldots x_{i-1} x_{i+1} \ldots x_m}}{1 - R^2_{yx_1 \ldots x_i \ldots x_m}} \cdot \frac{n-m-1}{1},$$

мұндағы $R^2_{yx_1 \ldots x_i \ldots x_m}$ – толық тізімді факторлары б модель үшін Жиынтық детерминация коэффициенті, $R^2_{yx_1 \ldots x_{i-1} x_{i+1} \ldots x_m}$ – дәл сол көрсеткіш, бірақ модельге x_i факторы енгізілмеген, n – зерттеулер саны, m – модельдегі параметрлер саны (бос мүшесіз).

Жеке *F*-критерийдің іс жүзіндегі мәні мәнділік деңгейі α болатын және еркіндік дәрежесінің саны 1 және n-m-1 болатын кестелік мәнімен салыстырылады. Егер F_{x_i} іс жүзіндегі мәні $F_{кесте}(\alpha, k_1, k_2)$ асса, онда модельге қосымша x_i факторын енгізу статистикалық расталады және таза регрессияның b_i коэффициентінің x_i факторы кезінде статистикалық мәні бар. Егер F_{x_i} іс-жүзіндегі мәні $F_{кесте}(\alpha, k_1, k_2)$аз болса, онда модельге қосымша x_i факторын енгізу y мәнінің түсіндірілген

вариациясының үлесін көбейтпейді, бұдан шығатыны, оны модельге енгізу дұрыс емес; берілген фактор кезінде регрессия коэффициентінің статистикалық мәні жоқ. Екіфакторлы теңдеу үшін жеке *F*-критерийлер түрі келесідей болады:

$$F_{x_1} = \frac{R^2_{yx_1x_2} - r^2_{yx_2}}{1 - R^2_{yx_1x_2}} \cdot (n-3), F_{x_2} = \frac{R^2_{yx_1x_2} - r^2_{yx_1}}{1 - R^2_{yx_1x_2}} \cdot (n-3)$$

Жеке *F*-критерийдің көмегімен Жиынтық регрессия теңдеуіне тиісті x_i факторы ең соңғы болып енгізілді деген тұжырымындағы барлық регрессия коэффициенттерінің мәнділігін бағалауға болады.

Жеке *F*-критерий таза регрессия коэффициенттерінің мәнділігін бағалайды. F_{x_i} көрсеткішін біліп, *i*-інші фактор кезінде регрессия коэффициенті үшін *t*-критерийді анықтауға болады, ал дәлірек:

$$t_{b_i} = \sqrt{F_{x_i}}.$$

Стьюденттің *t*-критерийі бойынша таза регрессияның коэффициенттерінің мәнділігін бағалануы жеке *F*-критерийді есептемей-ақ жасалады. Бұл жағдайда, жұптық регрессиядағыдай, әрбір фактор үшін келесі формула қолданылады:

$$t_{b_i} = \frac{b_i}{m_{b_i}},$$

мұндағы b_i – x_i факторы кезінде таза регрессия коэффициенті; m_{b_i} – регрессияның b_i коэффициентінің орташа квадраттық (стандарттық) қатесі.

$\hat{y} = a + b_1x_1 + b_2x_2 + ... + b_mx_m$ Жиынтық регрессия теңдеуі үшін регрессияның коэффициентінің орташа квадраттық қатесі келесі формуламен анықталады:

$$m_{b_i} = \frac{\sigma_y \sqrt{1 - R^2_{yx_1...x_m}}}{\sigma_{x_i} \sqrt{1 - R^2_{x_ix_1...x_m}}} \cdot \frac{1}{\sqrt{n-m-1}},$$

мұндағы σ_y – y мәні үшін орташа квадраттық ауытқу, σ_{x_i} – x_i мәні үшін орташа квадраттық ауытқу, $R^2_{yx_1\ldots x_m}$ – Жиынтық регрессия теңдеуі үшін детерминация коэффициенті, $R^2_{x_i x_1\ldots x_m}$ – Жиынтық регрессия теңдеуінің x_i факторының барлық басқа факторларымен тәуелділік кезіндегі детерминация коэффициенті; n-m-1 – ауытқу квадраттарының қалдық сомасы үшін еркіндік дәрежесінің саны.

Берілген формуланы қолдану үшін фактораралық корреляцияның матрицасы қажет және ол бойынша $R^2_{x_i x_1\ldots x_m}$ детерминация коэффициенттерін есептеу керек. Осылайша, $\hat{y} = a + b_1x_1 + b_2x_2 + b_3x_3$ теңдеуі үшін b_1, b_2, b_3 регрессияның мәнділігін бағалау үш $R^2_{x_1x_2x_3}$, $R^2_{x_2x_1x_3}$, $R^2_{x_3x_1x_2}$, фактораралық детерминация коэффициенттерін есептеулерін көздейді.

Таза регрессия коэффициенттері үшін корреляцияның жеке коэффициентінің, жеке F-критерийдің және Стьюденттің t-критерийінің көрсеткіштерінің өзарабайланысы факторларды таңдау процедурасында қолданыла алады.

Регрессия теңдеуін құру кезінде жою әдісі арқылы факторларды сүзуді әрбір қадамда корреляцияның жеке коэффициентінің ең мәнсіз мәнін жоя отырып, корреляцияның жеке коэффициенттерімен ғана емес, t_{b_i} және F_{x_i} көрсеткіштері бойынша жүзеге асыруға болады. Жеке F-критерий айнымалыларды енгізу әдісімен және қадамдық регрессиялық әдіспен модельді құру кезінде кең қолданылады.

1.27-тапсырма. Тексерілген нәтижелер бойынша 20 үй шаруашылығы ең кіші квадраттар әдісі бойынша тұрғызылған регрессия теңдеуі үшін тұтынатын шығындар адам басына шаққанда

$$Y = 120 + 0.62W + 0.36S.$$
$$(54) \quad (0.23) \quad (0.14)$$

Жақшада тиісті стандартты қателері көрсетілген. Маңыздылық коэффициентін тексеріңіздер. Коэффициенттік модель үшін 95 %-дық сенімді интервалын табыңыздар [8,10].

1.28-тапсырма. 1.28-кестеде x_1, x_2 айналымдары мен *y* мақсат функциясының мәндерін түсіндіретін T=10 ақпараты берілген.

1.28-кесте

x_{1t}	10,3	18,5	16,3	22,5	10,5	16,8	14,0	19,1	13,0	18,0
x_{2t}	2,5	8,6	3,7	6,5	7,8	9,1	1,9	2,7	3,0	5,2
y_t	24,8	48,3	37,0	51,8	29,1	43,0	30,1	41,0	29,1	40,1

Анықтау керек:

1) Сызықтық екі факторлық регрессияның теңдеуінің параметрлерін ЕКӘ-ң көмегімен бағалау керек $y_1 = \alpha_0 + x_{1t} + \alpha_2 x_{2t} + \varepsilon_1$;

2) R^2 детерминацияның коэффициентін мағынасын есептеп түсіндіру.

Бақылау сұрақтары

1. Корреляцияның жеке коэффициенттері қалай есептеледі?
2. Жиынтық регрессия теңдеуі үшін Фишердің F-критерийі және Фишердің жеке F-критерийі қандай болады?
3. Жиынтық регрессия теңдеуі үшін Стьюденттің t-критерийі қандай болады.

2. ЛОГИСТИКАЛЫҚ ЖҮЙЕЛЕРДЕГІ БОЛЖАУ ӘДІСТЕРІ

2.1. Болжау жасау теориясының негізгі ережелері

Қамсыздандыру, өндірістік, тарату логистикасында болжау жасаудың әдістері қолданылады, өйткені талданатын үдерістің дамуын болжаулық бағалары оперативтік, тактикалық және стратегиялық жоспарлау кезінде басқарушылық шешім қабылдаудың негізі болып табылады. Әртүрлі логистикалық операциялар мен функциялардың жүзеге асыру тиімділігі болжаудың дәлдігі мен кепілділігіне байланысты [27,29,31,34].

Кездейсоқ әсер ету болған кезде мәліметтерді талдаумен байланысты осындай есептерді шешу үшін болжау жасаудың статистикалық әдістерінің оның ішінде қолданбалы статистиканың маңызы өте зор. Оның әдістер кездейсоқтық заңдылықтарды айқындауға, негізделген болжау жасауға және олардың орындалу ықтималдығын бағалауға мүмкіндік береді.

Болжау болашақта әйтеуір бір құбылыстың мүмкін күйлері, келешегі немесе оларды жүзеге асыру мерзімі мен баламалық жолдары туралы ықтималды ғылыми негізделген пікір ретінде анықталады. *Болжау жасау әдістемесі* ретінде болжау жасау жүйесі және әдістері, тәсілдері туралы білімдер облысы түсіндіріледі, дәлірек айтқанда:

– болжау жасау әдісі – болжауының орындалуына негізделген объекті зерттеу тәсілі;

– болжау жасау методикасы – бір немесе бірнеше әдістердің жиынтығы;

– болжау жасау жүйесі – түрлі болжау әдістерін жүзеге асыру құралдарының реттелген жиыны.

Болжау жасау теориясына келесілер кіреді: болжау жасау объектісін талдау; математикалық және сараптамалық болып бөлінетін болжау жасау әдістері; болжау жасау жүйелері, оның ішінде үздіксіз жүйелер, мұнда мониторинг есебімен объектінің қызмет ету үдерісін тікелей жүзеге асырыу болып табылады.

Болжау жасау теориясының негізіне сүйене отырып объектілерді талдау кезінде болжаулардың классификациясы

жүргізіледі, оның негізгі белгілерін төмендегідей қарастыруға болады:

– объекті сипаттау кезінде мәнді айнымалылардың мөлшерін көрсету;

– айнымалылардың өзара байланыс деңгейін сипаттайтын күрделілігі;

– айнымалылардың детерминациялануы және стохастикалылығы;

– толық мөлшерлік қамтамасыз етілуі бар объектілерден қамтамасыз етілуі жоқ объектілерге барлық мүмкін нұсқалар енгізілген ретроспекция кезеңінің ақпараттық қамтамасыз етілуі.

Негізгі классификациялық қасиеттердің бірі болжау кезеңі болып табылады, мұнда болжаудың үш түрі ажыратылады: қысқа мерзімді (1 жылдан аспайды), орта мерзімді (1 жылдан 5 жылға дейінгі мерзім) және ұзақ мерзімді (5 жылдан астам).

Тек болжау ғана емес, болжау жасау әдістері де классифицияланылады. Болжау жасау әдістерінің мөлшері көп, бірақ олар екі тәсілдемеге негізделген:

– эвристикалық;

– математикалық.

Эвристикалық әдістер.

Оған ұсынылған ақпарат, тәжірибе, сезгіштік және сараптаушы білімі негізінде сараптаушылармен қалыптастырылған интуитивтік болжаудың модельдерін құру кіреді.

Мұнда келесі модельдер қарастырылады:

– жеке (сұхбат, ой генерациясы типіндегі модельдер);

– ұжымдық (қарапайым түзету әдісі; салмақтық коэффициенттерді беру әдісі; кезектілік салыстырулар әдісі; жұптық салыстырулар әдісі);

– комбинациялық («Дельфи» әдісі және оның модификациялары) сараптау бағалау.

Математикалық әдістер.

Болжау жасаудың математикалық әдістері үш топқа бөлінеді:

– уақыттық қатарлар бойынша экстраполяциялық симплекстік әдістері (ең кіші квадраттар әдісі, экспоненциалдық тегістеу, т.б.);

– корреляциялық және регрессиялық талдау, факторлық талдау кіретін статистикалық әдістер;

– болжаудың әртүрлі типтерінің синтезін құрайтын комбинациялық әдістер.

I типтегі болжаулар

I типтегі болжаулар *уақыттық қатарлар* негізінде симплекстік немесе статистикалық әдістерді қолдану арқылы жүзеге асырылады. I типтегі болжаулар келесі мәліметтерге тән:

– мәні бар айнымалылар саны 1-ден 3-ке параметр кіреді, өйткені ол масштабтылығы бойынша сублокальді болжауларға жатады;

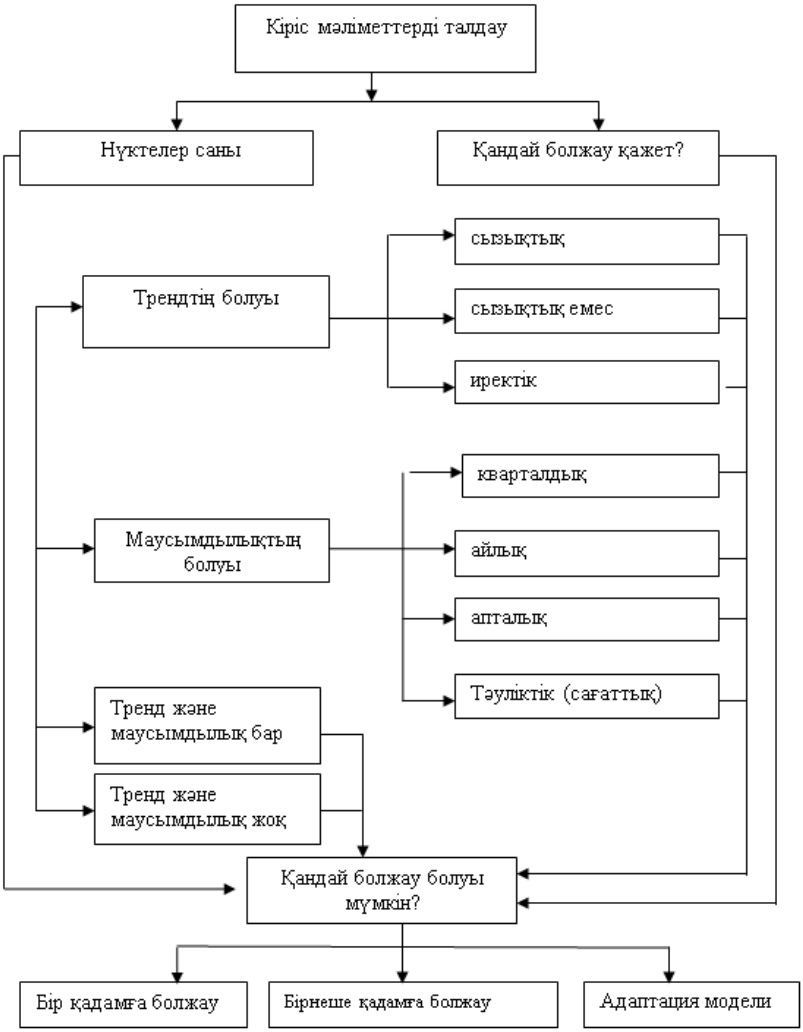
– бір – параметрлі, мысалы уақытты қолданған кезде бұл болжаулар өте қарапайым, ал екі – үш өзара байланысты параметрлер кезінде күрделі болып саналады;

– I типтегі болжаулар болашақ кезеңінің ақпараттық қамсыздандыру деңгейі бойынша толық ақпараттық қамтамасыз етілуі бар объектілерге жатады.

II типтегі болжаулар.

Негізінен, II типтегі болжаулар ұзақ мерзімді болжау жасау үшін қолданылады және екі кезеңге бөлінеді: *біріншісі* – негізгі факторлардың болжаулық бағаларын алу; екіншісі – үрдістің немесе құбылыстың даму болжауы.

Болжау құру тиімді болу үшін болжаулық бағаларды есептеу тәсілдері туралы білім жеткіліксіз. Болжау жасаудың әртүрлі әдістерін пайдалану облыстарының шекараларын білу қажет және мәліметтердің көлемі мен сипаттарына байланысты нақты болжау алу мақсатына сәйкес қажетті әдісті жылдам және қатесіз таңдау керек.



2.1-сурет. Нақты болжау алу үшін болжау мақсаттарын, алғашқы мәліметтер көлемін және мәліметтер түрін сәйкестендіруге мүмкіндік беретін сызба

Болжау жасағанда оны таңдаудың жалпы ұсыныстары*:*

Болжау жасау әдісі	Трендтің болуы	Маусым-дылықтың болуы	Мәліметтердің минималды мөлшеріне қойылатын талаптар		Болжау жасау мүмкіндігі	
			Маусымдық емес	Маусымдық	Бір қадам	Бірнеше қадам
Сараптық	Міндетті емес	Міндетті емес	0	0	Ия	Ия
Аңғырт	Ия /Жоқ	Ия/жоқ	1	-****	Ия	Жоқ
Экспоненциалдық тегістеу (қарапайым)	Жоқ	Жоқ	2	-	Ия	Жоқ
Арифметикалық тегістеу	Жоқ	Жоқ	4	-	Ия	Жоқ
Хольт әдісі	Ия	Жоқ	3	-	Ия	Жоқ
			10	-	Ия	Ия
Экспоненциалдық орташа әдісі (Браун әдісі)	Ия	Жоқ	3	-	Ия	Ия *
Трендтің экстраполяциясы	Ия	Жоқ	3	-	Ия	Ия*
	Ия	Ия	-	1 х Т**	Ия	Ия *
Винтерс әдісі	Ия	Ия	-	2 хТ	Ия ***	
Комбинациялық болжау (орташа мәнді бағалау)	Ия/жоқ	Ия/жоқ	1	-	Ия	Жоқ
Комбинациялық болжау (орташа мәнді және ауытқуды бағалау)	Ия/жоқ	Ия/жоқ	3	-	Ия	Жоқ
*бірнеше қадамға болжау жасау болжау жасау алдындағы кезең және болжау жасау кезеңінің ұзындықтарының қатынасын 3:1 сақтау кезінде мүмкін. **Т – маусымдылықтың кезеңділігі. ***Маусымдылығы бар бір кезеңге болжау (мысалы, бір жылға квартал сайын). **** «-» – маусымдылықты есептеу үшін әдіс қолданылмайды.						

І типтегі болжау жасау үшін уақыттық қатарларды талдауға негізделген әдістер жиі қолданылады. Бұл әдістердің күрделілігі әртүрлі.

Осылайша, мысалы, егер қолдағы барлық мәліметтер мәнді және олардың құндылығы болжау үшін бірдей болса, динамикалық қатардың барлық мәндерінің орташасын табуға болады, бұл бір қадамға жасалған болжау болады. Бұл әдісті мәліметтердің кез келген мөлшері кезінде барлық мәліметтердің құндылығы тең деген шарт қолдануға болады. Бірақ егер динамикалық қатардың соңғы мәні басқалармен салыстырғанда мәндірек болса, онда ол бір қадамға жасалынған болжау ретінде

алына алады. Бұл болжау жасаудың қарапайым әдістері іс жүзінде ешқандай есептеулерді қажет етпейді.

Бақылау сұрақтары

1. Болжау дегеніміз не?
2. Болжаудың қандай түрлері бар?
3. Логистикадағы болжаудың түрлерін атаңыз.

2.2. Уақыттық қатарлар

Белгілі уақытта логистикалық үрдістің дамуы статистикалық сипаттамасы бар уақыттық қатарлар көмегімен жүзеге асырылады. Уақыттық қатар деп хронологиялық кезектілікпен, яғни t-уақыттық параметрі айнымалысының өсу ретімен берілген кез келген көрсеткіштің мәндеріне жасалған бірқатар зерттеулерді айтамыз. Уақыттық қатардың арнайы зерттеулері бұл қатардың деңгейлері деп аталады [3,18,19,20]. Мезеттік уақыттық қатарларда деңгейлер уақыттың белгілі мезеттеріндегі күйінің мәндерін сипаттайды.

Мысалы, тауарлардың белгілі түрлерінің бағаларының уақыттық қатарлары, акциялар курстарының уақыттық қатарлары мезеттік болып табылады.

Интервалдық қатарларда деңгейлер уақыттың белгілі интервалдағы (кезеңдерінде) көрсеткіштің мәнін сипаттайды. Бұл типтегі қатарлардың мысалдары бір ай, квартал, жыл, т.б. натуралдық немесе құндық өрнектермен берілген өнім өндірудің уақыттық қатарлары болады.

Кейде қатар деңгейлері зерттелетін мәндерді, ал туынды шамалар орташа және қатысты шамаларды құрайды. Мұндай қатарлар туынды деп аталады. Мұндай уақыттық қатарлардың деңгейлері зерттелетін көрсеткіштер негізінде кейбір есептеулер көмегімен алынады.

Мұндай қатарлардың мысалдары ретінде өндірістік өнімнің негізгі түрлерін орташа тәуліктік өндіріс қатарлары және бағалар индекстерінің қатарлары қарастырылады.

Қатар деңгейлері детерминацияланған немесе кездейсоқ мәндер қабылдай алады.

Айлардағы күндер саны туралы мәліметтер қатары деңгейлердің детерминацияланған мәндері бар қатардың мысалы болады. Талдауға, кейінірек, болжау жасауға деңгейлердің кездейсоқ мәндері бар қатарлар да түседі. Мұндай қатарларда әрбір деңгей дискретті немесе үздіксіз кездейсоқ шаманың жүзеге асырылуы ретінде қарастырылуы мүмкін.

2.1-кестеде уақыттық қатарлардың мысалдары келтірілген: бірінші қатар мезеттік, екінші қатар – интервалдық. Үшінші уақыттық қатардың деңгейлері – есептелген шамалар, ал айлық динамиканың уақыттық қатары туынды болып табылады.

2.1-кесте

Уақыттық қатарлар мысалдары

II) Кәсіпорын жұмысшыларының төлем ақы қоры (мың тг)			III) Кәсіпорындағы орташа тәуліктік өнім өндіру (дана)	
Ай	t	y_t	t	y_t
Қаңтар	1	79,5	1	1570
Ақпан	2	84.1	2	1590
Наурыз	3	85,5	3	1595
Сәуір	4	88,5	4	1603
Мамыр	5	89,9	5	1610
Маусым	6	90,0	6	1600

Сауда жабылу мезетіндегі өндірістік компанияның акциялар бағасы (тг)		
Мерзімі	T	y_t
6.9.99	1	383
7.9.99	2	392
8. 9.99	3	391
9. 9.99	4	399
10. 9.99	5	397
13. 9.99	6	399

Үрдісті ары қарай зерттеу үшін қатардың көрші деңгейлері арасында интервалдарды таңдау маңызды орын алады. Бір біріне тең орналасқан қатар деңгейлерімен жұмыс істеген ыңғайлырақ. Мұнда, егер уақыттың өте үлкен интервалы таңдалса, көрсеткіш динамикасындағы маңызды заңдылықтары қаралмай қалуы мүмкін.

Әрине, қатардың деңгейлері арасында уақыт интервалын таңдау туралы мәселе әрбір нақты зерттеудің мақсаттарына сәйкес шешілуі тиіс.

Уақыттық қатарларды болжам жасау үрдісі бұдан бұрынғы үрдістің динамикасын түсіндіретін заңдылықтарды айқындауға және бұл заңдылықтарды келешекте дамуды сипаттау үшін қолдануға негізделген.

Мұнда даму және болжам жасауға талдау жасау, негізінен, кіріс ақпаратқа белгілі талаптар қоятын математикалық аппаратқа сүйенеді.

Мысалы, кішігірім кәсіпорындар санын сипаттайтын уақыттық қатардың деңгейлері «кішігірім кәсіпорын» деген түсініктің өзгерісіне байланысты сәйкес емес болып қалуы мүмкін. Бұл түсінік бүкіл зерттеліп отырған кезең үшін бірдей болуы керек. Көбінесе сәйкессіздік құндық көрсеткіштерде кездеседі, бұл талданатын кезеңде бағалар өзгерісімен түсіндіріледі.

Сәйкессіздік территориалдық өзгерістер нәтижесінде пайда болуы мүмкін.

Мысалы, мемлекеттің, аймақтың, облыстың шекарасының өзгеру нәтижесі ретінде. Сәйкессіздіктің басқа себебі құрылымдық өзгерістер болады, мысалы, бір бүтінге қосылу есесінен бірнеше ведомстволардың үлкеюі, немесе бірнеше кәсіпорындардың бір бірлестікке бірігуі есесінен өндірістің үлкеюі.

Белгілі математикалық аппаратты қолдану уақыттық қатарлардың мүмкін ұзындығына шектеу қояды. Мысалы, регрессиялық талдау қолдану үшін ұзындығы тәуелсіз айнымалылардың мөлшерін бірнеше рет артатын уақыттық қатарлар болуы керек.

Уақыттық қатарлардың деңгейлерінің аномальді мәндері немесе «шығарылу» болуы мүмкін. Көбінесе ақпаратты жинау, жазу немесе жіберу кезінде болатын қателермен түсіндіріледі. Қате мәндердің пайда болудың мүмкін көздері: құжаттан ақпаратты ауыстыру кезінде үтірдің жылжуы, мәліметтерді басқа графаға жазу, т.б.

Мұндай мәндерді айқындау, алып тастау, оларды ақиқат немесе есептеу мәндерімен алмастыру мәліметтерді алғашқы өңдеудің қажетті кезеңі болып табылады, өйткені «ластанған» ақпаратқа математикалық әдістерді қолдану талдау нәтижелерінің өзгеруіне алып келеді. Бірақ, аномальді мәндер үдерістің нақты дамуын, мысалы, «қара сейсенбідегі» доллар курсының «секіруін» көрсетуі мүмкін.

Негізінен, бұл мәндер модельдерді құру кезінде есептеу мәндерімен алмастырылады, бірақ модель бойынша алынған нақты мәндердің ауытқуының мүмкін шамасын есептеу кезінде ескеріледі.

Кіріс ақпараттың барлық көрсетілген талаптарға сәйкестігі уақыттық қатарларды алдын ала талдау кезеңінде тексеріледі. Осыдан кейін ғана даму динамикасының негізгі көрсеткіштерін есептеуге және талдауға, болжам жасау модельдерін құруға, болжамдық бағаларды алуға кіріседі.

Бақылау сұрақтары

1. Уақыттық қатарлардың қандай түрлерін білесіз? Мысалдар келтіріңдер.
2. Уақыттық қатарлардың кеңістіктік іріктемелерден қандай айырмашылығы бар екенін түсіндіріңіз?
3. Болжау кезінде уақыттық қатарларға кіріс ақпараты ретінде қандай талаптар қойылады?

2.3. Уақыттық қатарлардың компоненттері. Тенденцияның бар болуы туралы гипотезаны тексеру

Болжам жасау тәжірибесінде логистикалық көрсеткіштердің уақыттық қатарлар деңгейлерінің мәндері келесі компоненттерден тұрады деп есептелген: *тренд, маусымдық, циклдық және кездейсоқ құрамалар*.

Тренд деп уақыттық қатардың негізгі тенденциясын, дамудың жалпы бағытын анықтайтын өзгерісін айтады. Бұл ұзақ мерзімді әрекеттің жүйелік құрамасы [20,22,23].

Ұзақ мерзімді тенденциялармен қатар логистикалық үрдістердің уақыттық қатарларында жүйелі тербелістер динамиканың қатарлардың периодтық құрамалар орын алады.

Егер тербелістер периоды 1 жылдан аспаса, оларды маусымдық деп атайды. Көбінесе оларды пайда болу себебі табиғи – климаттық шарттар. Кейде маусымдық тербелістер себептеріне факторлар жиынтығына тәуелді сатып алу қабілеттілігі немесе халық сұранысының маусымдық сипаты жатады.

Мысалы, мереке алдындағы сатып алулардың көбеюі, квартал аяғындағы төлемдердің көбеюі, т.б.

Тербеліс периоды үлкен болғанда, уақыттық қатарларда *циклдік құрама* орын алады. Оның мысалдары демографиялық, инвестициялық және басқа циклдер.

Егер уақыттық қатардан трендті және периодтық құрамаларды алып тастасақ, тұрақты емес компонента қалады.

Экономистер тұрақты емес компонента қалыптастырылатын факторларды 2-ге бөліп қарастырады:

– бірден, кенеттен жасалынған іс-әрекет нәтижесінде пайда болған факторлар;

– ағымдағы факторлар.

Факторлардың бірінші типі (мысалы, келеңсіз оқиғалар, эпидемиялар, т.б.), негізінен, олар кездейсоқ тербелістермен салыстырғанда көбірек ауытқулар туғызады, мұндай ауытқуларды катастрофалық тербелістер деп атайды.

Факторлардың екінші типі үстеме себептердің үлкен санының әсер ету нәтижесі болып табылатын кездейсоқ

тербелістерді туғызады. Ағымдағы факторлардың әрқайсысының әсері аз, бірақ олардың сомалық әсер етуі байқалады.

Егер уақыттық қатар тиісті компоненталардың сомасы түрінде берілсе, онда алынған логистикалық модель аддитивтік (2.1) деп аталады, егер көбейтінді түрінде – мультипликативтік (2.2) немесе аралас типте болады (2.3):

$$Y_t = u_t + s_t + v_t + e_t, \quad (2.1)$$

$$Y_t = u_t \cdot s_t \cdot v_t \cdot e_t, \quad (2.2)$$

$$Y_t = u_t \cdot s_t \cdot v_t + e_t, \quad (2.3)$$

мұндағы Y_t – уақыттық қатарлар деңгейлері;

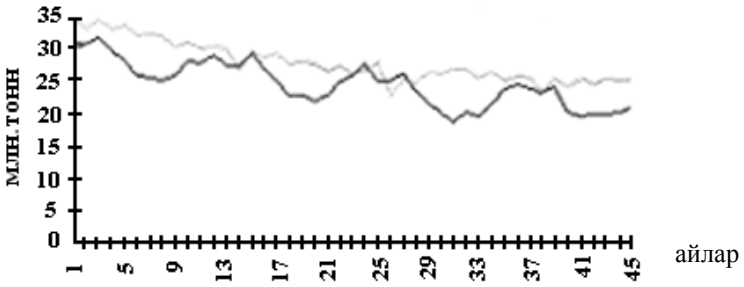
u_t – трендтік құрама;

s_t – маусымдық құрама;

v_t – циклдік құрама;

e_t – кездейсоқ компонента.

2.2–2.3 суреттерде көрсетілген компонеттердің болуын көрсететін уақыттық қатарлардың мысалдары келтірілген. Өндірістік өнімді өндірудің айлық уақыттық қатарлардың графиктері төмендеп келе жатқан тренд кезінде тұрақты маусымдық тербелістерді көрсетеді, мұнда соңғы учаскеде өндірістің түсу қарқыны төмендейді.

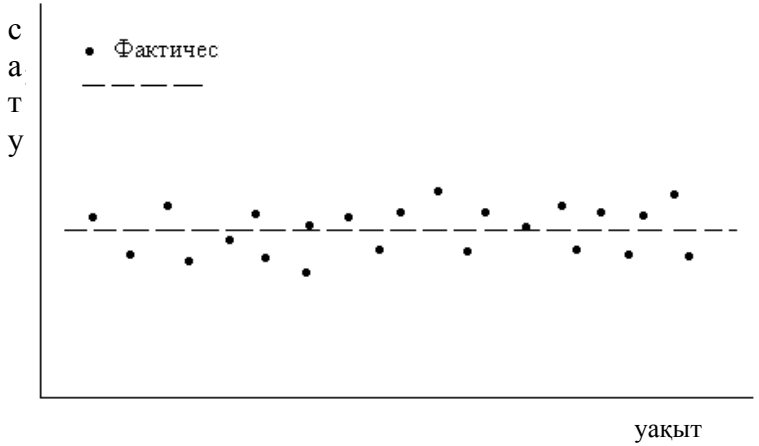


2.2-сурет. Натуралды өрнекпен берілген өндірістік өнімнің арнайы түрлерін өндірудің айлық динамикасы

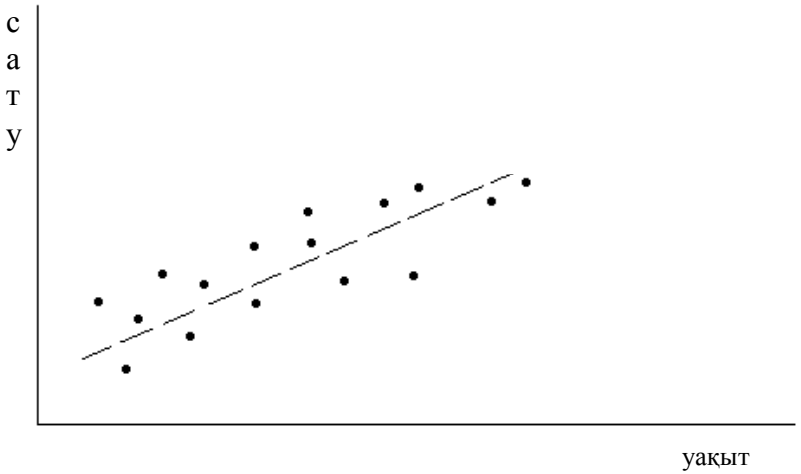
Уақыттық қатарларды болжау жасау және талдау бойынша кез келген есептің шешімі зерттелетін логистикалық көрсеткіштердің графигін тұрғызудан басталады. Мұнда уақыттық қатардағы трендтің болуы анық болып үнемі көрінбейді. Мұндай жағдайларда, тенденцияны анықтау және трендті көрсету алдында, зерттелетін үдерісте тенденцияның бар жоғын анықтау керек.

Сұраныс түрлерін көрсететін уақыттық қатарлардың мысалдарын келтіреміз (2.3,а–2.3,с-суреттер).

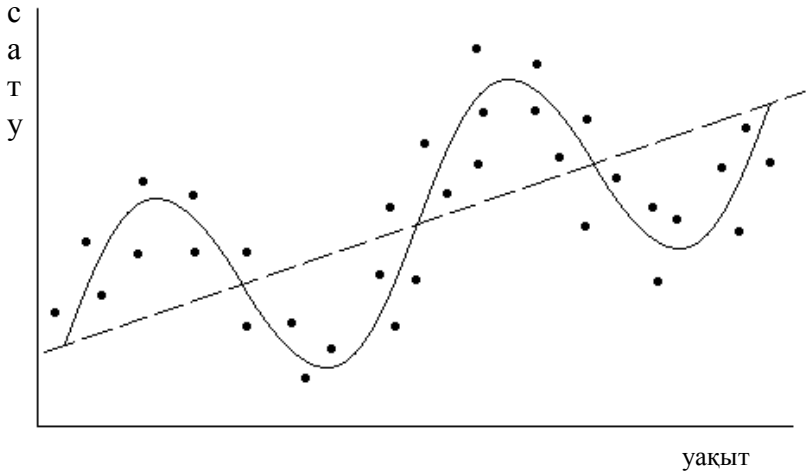
Кейбір жағдайларда тұрақты емес сұранысы бар тауарлар сатылатын тауарлардың жалпы көлемінің 50 % құрайды. Бұл жағдайда сатылымдардың болжамын жасау логистика жүйесі үшін қиын мәселе болып табылады.



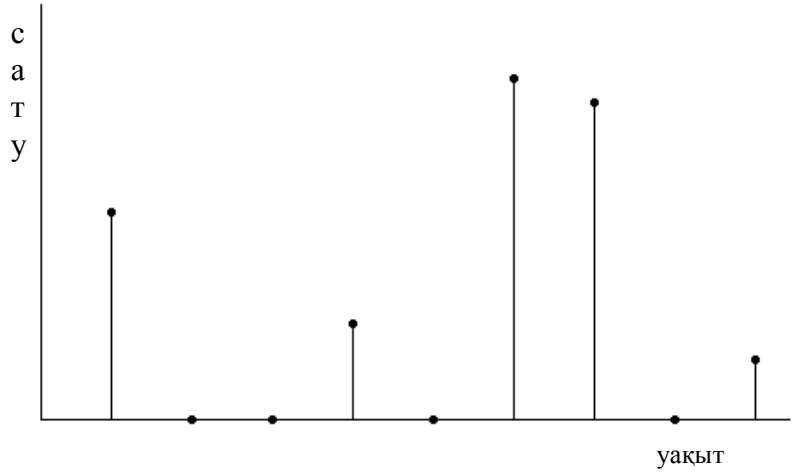
2.3,а-сурет. Тербелістердің тұрақты орташа деңгейі бар жүйелі сұраныс



2.3,b-сурет. Өспелі тренді бар жүйелі сұраныс



2.3,c-сурет. Өспелі тренді және маусымдылығы
бар жүйелі сұраныс



2.4-сурет. Тұрақты емес сұраныс

Бақылау сұрақтары

1. Уақыттық қатарлар деңгейлерінің мәндері қандай компоненттерден тұрады?
2. Трендке анықтама беріңіз.
3. Маусымдық компонентаның пайда болу себептерін келтіріңіз.
4. Циклдік және кездейсоқ компоненталарға мысалдар келтіріңіз.

2.4 Динамиканың негізгі көрсеткіштері және уақыттық қатарларды тегістеу үшін жылжымалы орташаларды қолдану

Уақыттық қатарлар динамикасын мөлшерлік бағалау үшін келесі статистикалық көрсеткіштер қолданылады: абсолюттік өсім, өсу қарқыны, өсім қарқыны, олар тізбекті, базистік және орташа болып бөлінеді[18,31].

Динамиканың бұл көрсеткіштерін есептеу негізінде уақыттық қатардың деңгейлерін салыстыру жатыр. Егер

салыстыру базасы ретінде алынған бір деңгеймен жүзеге асырылса, бұл көрсеткіштер *базисті* деп аталады. Егер салыстыру айнымалы база кезінде жүзеге асырылса, онда осылайша есептелген көрсеткіштер *тізбекті* деп аталады.

Абсолюттік өсім Δ_y салыстырылатын екі деңгейлердің айырмасына тең $\Delta y_t = y_t - y_{t-1}$.

Өсу қарқыны пайызбен берілген қатардың салыстырылатын екі деңгейінің қатынасын сипаттайды. $T_t = \frac{y_t}{y_{t-1}} \times 100°/\circ$

Өсім қарқыны қатысты шамалардағы абсолюттік өсімді сипаттайды $K_t = T_t - 100°/\circ$. Пайызбен анықталған өсім қарқыны салыстырылатын деңгей салыстыру базасы деп қабылданған деңгейге қатысты қанша пайызға өзгергенін көрсетеді. 2.2-кестеде базистік және тізбекті өсімді, өсу қарқынын, өсім қарқынын есептеу үшін өрнектер келтірілген. Мұнда келесі белгілеулер пайдаланылған: $y_1, y_2, ..., y_t, ..., y_n$

t – уақыттық қатарлар деңгейлері t=1, 2, ... , n;

n – уақыттық қатарлардың ұзындығы (3, 5, 10 немесе 15 жыл);

$y_б$ – салыстыру базасы деп қабылданған уақыттық қатардың деңгейі.

2.2-кесте

Динамиканың негізгі көрсеткіштері

	Өсім	Өсу қарқыны	Өсім қарқыны
Тізбекті	$\Delta y_t = y_t - y_{t-1}$	$T_t = \frac{y_t}{y_{t-1}} \times 100°/\circ$	$K_t = T_t - 100°/\circ$
Базистік	$\Delta y_б = y_t - y_б$	$T_б = \frac{y_t}{y_б} \times 100°/\circ$	$K_{б_t} = T_{б_t} - 100°/\circ$
Орташа	$\overline{\Delta \bar{y}} = \frac{y_n - y_1}{n-1}$	$\overline{T} = \sqrt[n-1]{\frac{y_n}{y_1}} \times 100°/\circ$	$\overline{K} = \overline{T} - 100°/\circ$

Даму динамикасының жалпылайтын көрсеткіштерін алу үшін орташа шамалар анықталады: орташа абсолюттік өсім, өсу мен өсімнің орташа қарқыны.

Орташа өсім көмегімен қатардың динамикасын сипаттау екі шекті нүкте арқылы жүргізілген түзу түріндегі көрінісіне сәйкес келеді. Бұл жағдайда бір қадам бұрын болжам алу үшін, соңғы зерттеуге орташа абсолюттік өсімнің мәнін қосу жеткілікті.

$$\hat{y}_{n+1} = y_n + \overline{\Delta y} \quad , \qquad (2.4)$$

мұндағы y_n – қатардың соңғы n-інші нүктесіндегі нақты мәні;

$\hat{y}_{n+1}$ – n+1 нүктесіндегі деңгей мәнінің болжамдық бағасы;

$\overline{\Delta y}$ – $y_1, y_2, ..., y_n$ уақыттық қатар үшін есептелген орташа өсім мәні.

Болжамдық мәнді алудың мұндай тәсілдемесі орынды, егер оның даму сипаты сызықтыққа жақын болса. Дамудың мұндай тең өлшемді сипатына тізбекті абсолюттік өсімнің бірдей мәндері сәйкес болады.

Қатардың динамикасын сипаттау үшін өсуді орташа қарқынын (өсімнің орташа қарқынын) қолдану екі шекті нүкте арқылы жүргізілген көрсетілген немесе экспоненциалды қисық түріндегі көрінісіне сәйкес келеді. Сондықтан бұл көрсеткішті жалпы түрде қолдану динамика өзгерісі өсудің тұрақты қарқынымен жүретін үдерістер үшін дұрыс болады. Бұл жағдайда і қадамға ерте жасалған болжамның мәні келесі формуламен алынады:

$$\hat{y}_{n+i} = y_n \times \overline{T}^i \, , \qquad (2.5)$$

$\hat{y}_{n+i}$ – $n+i$ нүктесіндегі деңгейдің мәндерін болжаудың бағасы;

y_n – қатардың соңғы n-інші нүктесіндегі нақты мәні;

$\overline{T}$ – $y_1, y_2, ..., y_n$ қатары үшін есептелген өсудің орташа қарқыны (%-берілмейді).

Орташа өсімнің және өсудің орташа қарқынының кемшіліктеріне олардың қатардың ақырғы және бастапқы

деңгейлерін ескеретіндігі жатады. Бірақ бұл көрсеткіштердің қолдану облысы кең, бұл оның есептелу қарапайымдылығымен түсіндіріледі. Олар терең сапалық және мөлшерлік талдауды қажет ететін болжамдаудың қарапайым әдістері ретінде қолданылуы мүмкін.

2.1-мысал. Біреуімен келісім-шартты ұзарту туралы шешім қабылдау үшін жұмыс нәтижелері бойынша № 1 және № 2 жеткізушілердің бағалауын жүргізу [17,26,28,29].

Есептеу мысалы. Алғашқы екі ай ішінде фирма № 1 және № 2 жеткізушілерден А және В тауарларын алады. Жеткізілетін осындай өнім бағасының динамикасы, сондай-ақ жеткізушілердің жеткізілімнің орнатылған мерзімдерін бұзу динамикасы 2.3–2.5 кестелерде келтірілген.

Жеткізушілердің біреуімен келісім-шартты ұзарту туралы шешім қабылдау үшін әрбір жеткізушінің рейтингін есептеу керек. Жеткізушілерді бағалауды келесі көрсеткіштер бойынша жасау керек: жеткізілімнің бағасы, кепілділігі, сапасы. Мұнда А және В тауарлары үздіксіз толықтыруды қажет етпейтінін ескеру керек. Осылайша, жеткізушілер рейтингін есептеу кезінде көрсеткіштердің келесі салмақтарын қабылдаймыз:

Бағасы	0,5
Жеткізілім бағасы	0,3
Жеткізілім сенімділігі	0,2

2.3-кесте

Жеткізілімдер бағасының динамикасы

Жеткізуші	Ай	Тауар	Жеткізілім көлемі, бірл/ай	Бірлігі үшін бағасы, тг
№ 1	қаңтар	А	2 300	11
	қаңтар	В	1 200	6
№ 2	қаңтар	А	9 300	10
	қаңтар	В	7 200	6
№ 1	ақпан	А	2 400	12
	ақпан	В	1 000	8
№ 2	ақпан	А	8 400	12
	ақпан	В	8 000	8

2.4-кесте

Қажетті емес сапада тауарларды жеткізу динамикасы

Ай	Жеткізуші	Бір ай ішінде қажетті емес сапада келген тауардың мөлшері, бірл.
Қаңтар	№ 1	90
	№ 2	70
Ақпан	№ 1	110
	№ 2	60

2.5-кесте

Жеткізілімнің орнатылған мерзімдерін бұзу динамикасы

Жеткізуші № 1			Жеткізуші № 1		
Ай	Жеткізілімдер саны, бірл.	Кешігулер саны,тәулік	Ай	Жеткізілімдер саны, бірл.	Кешігулер саны,тәулік
Қаңтар	9	20	Қаңтар	8	30
Ақпан	6	24	Ақпан	8	24

Жеткізушінің рейтингін есептеуді 2.8-кесте түрінде безендендіру керек.

1. Жеткізілімдер бағаларының өсу қарқынын (баға көрсеткіші) ($T_б$) есептеу

$$T_б = \sum T_{бi} d_i$$

мұндағы $T_{бi}$ – жеткізілімнің і-інші түрінің бағасының өсу қарқыны;

d_i – ағымдағы кезеңнің жеткізілімінің жалпы көлемінде тауардың і-інші түрінің үлесі;

n – жеткізілімдердің түрлерінің саны.

Жеткізілетін тауардың і-інші түрінің бағасының өсу қарқыны келесі формуламен есептеледі:

$$T_{бі} = (P_{i1} / P_{i0} 100),$$

мұндағы P_{i1} – ағымдағы кезеңнің тауардың і–інші түрінің бағасы;

P_{i0} – алдыңғы кезеңнің тауардың і–інші түрінің бағасы.

Жеткізілімінің жалпы көлемінде тауардың і-інші түрінің үлесі келесі формуламен есептеледі:

$$d_i = (S_i / \sum S_i),$$

мұндағы S_i – ағымдағы кезеңде тауардың і-інші түріне қойылған сомасы, сом.

Мысал ретінде бірінші жеткізуші үшін өсудің орташа өлшенген қарқынын есептейміз.

А тауары бойынша бұл жеткізуші үшін бағалардың өсу қарқыны:

$$T_{бА} = 12/11*100 = 109\%.$$

В тауары бойынша:

$$T_{бА} = 8/6x100 = 133\%.$$

Ағымдағы кезеңнің жеткізілімнің жалпы көлемінде А тауарының үлесі:

$$d_A = 2400 \cdot 12/(2400 \cdot 12 + 1000 \cdot 8) = 0{,}78,$$

ағымдағы кезеңнің жеткізілімнің жалпы көлемінде В тауарының үлесі:

$$d_в = 1000 \cdot 8/(2400 \cdot 12 + 1000 \cdot 8) = 0{,}22.$$

Бірінші жеткізуші үшін өсудің орташа өлшенген қарқыны:

$$T_б = 109 \cdot 0{,}78 + 133 \cdot 0{,}22 = 114{,}4\%.$$

Бағалардың өсуіндегі орташа өлшенген қарқынын есептеу үшін 2.6-кестені ұсынамыз.

2.6-кесте

Бағалардың өсуінің орташа өлшенген қарқынын есептеу

Жеткізуші	$Т_б$А	ТбВ	SA	SB	dA	dB	Т'б
№ 1	109	133	28 800	8 000	0,78	0,22	114,4
№2	120	133	100800	64 000	0,61	0,39	125,2

Жеткізушілердің рейтингін есептеу үшін алынған мәндер қорытынды 2.8-кестеге енгізіледі

2. Қажетті емес сапада тауарларды жеткізудің өсу қарқынын (сапа көрсеткішін) есептеу

Екінші көрсеткіш бойынша жеткізушілерді бағалау үшін (жеткізілім сапасы) әрбір жеткізуші бойынша қажетті емес сапада тауарларды жеткізудің өсу қарқынын есептейміз (*Тқ.е.с.*):

$$Тқ.е.с. = dқ.е.с.1 / dқ.е.с.0 \cdot 100,$$

мұндағы *dқ.е.с.1* – ағымдағы кезеңдегі жалпы көлеміндегі қажетті емес сападағы тауарлар;

dқ.е.с.0 – алдыңғы кезеңдегі жалпы көлеміндегі қажетті емес сападағы тауарлар.

Жеткізілімнің жалпы көлеміндегі қажетті емес сападағы тауарлардың үлесін 2.4-кесте және 2.5-кесте негізінде есептейміз. Нәтижелерді 2.8-кестеге жазамыз.

2.7-кесте

Жеткізілімнің жалпы көлеміндегі қажетті емес сападағы тауарлардың үлесін есептеу

Ай	Жеткізуші	Жалпы жеткізілім, бірл/ай	Жеткізілімнің жалпы көлеміндегі қажетті емес сападағы тауарлардың үлесі , %
Қаңтар	№ 1	3 500	2,6
	№ 2	16 500	0,4
Ақпан	№ 1	3 400	3,2
	№ 2	16 400	0,4

Біздің мысалымызда бірінші тұтынушы үшін қажетті емес сападағы тауарлардың өсу қарқыны: *Тқ.е.с. = 3,2 / 2,6 · 100 = 126%,*

Алынған нәтиже 2.8-кестеге енгізілген.

3. Орташа кешігудің өсу қарқынын есептеу (жеткізілім сенімділігінің көрсеткіштері).

Орташа кешігу жеткізілімнің сенімділігінің мөлшерлік бағасы болып қызмет етеді, яғни 1 жеткізілімге келген кешіккен күндер саны. Бұл шама белгілі кезеңде кешігу күндерінің жалпы санын сол кезеңдегі жеткізулер санына бөліндісі ретінде анықталады (2.5-кесте мәліметтері).

Осылайша, әрбір тұтынушы бойынша орташа кешігудің өсу қарқыны келесі формула бойынша есептеледі:

$$T_{орткеш.} = (O_{орткеш1} / O_{орткеш2}) \cdot 100 \ ,$$

мұндағы $O_{орткеш1}$ – ағымдағы кезеңнің бір ставкаға орташа кешігуі, күн;

$O_{орткеш.2}$ – алдыңғы кезеңде бір жеткізілімге орташа кешігуі, күн.

Ары қарай № 1 жеткізуші үшін орташа кешігудің өсу қарқынын есептейміз:

$$T_{ортакеш.} = (24:6)/(20:9) \cdot 100 = 180\%.$$

Алынған нәтижені 2.8-кестеге енгіземіз.

4. Жеткізушілердің рейтингін есептеу

Рейтингті есептеу үшін әрбір көрсеткіш бойынша өсу қарқынының өсу қарқыны бойынша салмаққа туынды алу керек. 2.8-кестенің 5 графасы бойынша көбейтінділер сомасы № 1 жеткізушінің рейтингін, ал 6 графа бойынша – № 2 жеткізушінің рейтингін береді.

Біздің жағдайда, өсу қарқыны жеткізушінің негативті сипаттамаларының (бағалардың өсуі, жеткізілімнің жалпы көлемінде сапалы емес тауарлардың үлесінің өсуі, кешігулер мөлшерінің өсуі) артуын көрсететіндігін ескере отырып, біз келісім-шартты қайта жасау кезінде берілген әдістеме бойынша есептелген рейтингі төмен жеткізушіні таңдауымыз керек.

2.8-кесте

Жеткізушілердің рейтингін есептеу

Көрсеткіштер	Көрсеткіш салмағы	Берілген көрсеткіш бойынша жеткізушіні бағалау		Салмаққа бағалардың көбейтінділері	
		Жеткізуші №1	Жеткізуші №2	Жеткізуші №1	Жеткізуші №2
Баға	0,5	114,4	125,2	57,2	62,6
Сапа	0,3	126	86	37,7	25,9
Сенімділігі	0,2	180	80	36,0	16,0
Жеткізуші рейтингі				130,9	104,5

2.1-тапсырма. Қазақстан Республикасы аймағының бірнеше жылдардағы өнім экспорты (млн.долл.) туралы ақпарат бар (2.9-кесте) [38]:

2.9-кесте

Жылдар	2013	2014	2015	2016
Өнім экспорты	84 700,4	79 459,8	45 955,8	36 775,6

Анықтау керек:

1) тізбектік және базистік:

а) абсолюттік өсімдерді; б) өсу қарқыны; в) өсім қарқыны;

2) өсімнің бір пайызының абсолюттік мазмұны;

3) орташа көрсеткіштер: а) қатардың орташа деңгейі;

б) орташа жылдық абсолюттік өсім; в) орташа жылдық өсім қарқыны.

2.2-тапсырма. Алғашқы екі ай ішінде фирма № 1 және № 2 жеткізушілерден А және В тауарларын алады. Жеткізілетін осындай өнім бағасының динамикасы, сондай-ақ жеткізушілердің жеткізілімнің орнатылған мерзімдерін бұзу динамикасы 2.10–2.12 кестелерде келтірілген.

Жеткізушілердің біреуімен келісім-шартты ұзарту туралы шешім қабылдау үшін әрбір жеткізушінің рейтингін есептеу керек. Жеткізушілерді бағалауды келесі көрсеткіштер бойынша

жасау керек: жеткізілімнің бағасы, кепілділігі, сапасы. Мұнда А және В тауарлары үздіксіз толықтыруды қажет етпейтінін ескеру керек. Осылайша, жеткізушілер рейтингін есептеу кезінде көрсеткіштердің келесі салмақтарын қабылдаймыз:

Бағасы	0,8
Жеткізілім бағасы	0,5
Жеткізілім сенімділігі	0,3

2.10-кесте

Жеткізілімдер бағасының динамикасы

Жеткізуші	Ай	Тауар	Жеткізілім көлемі, бірл/ай	Бірлігі үшін бағасы, тг
№1	Қаңтар	А	2 000	10
	Қаңтар	В	1 000	5
№2	Қаңтар	А	9 000	9
	Қаңтар	В	8000	5
№1	Ақпан	А	2 400	10
	Ақпан	В	1 000	6
№2	Ақпан	А	9300	12
	Ақпан	В	8 400	8

2.11-кесте

Қажетті емес сапада тауарларды жеткізу динамикасы

Ай	Жеткізуші	Бір ай ішінде қажетті емес сапада келген тауардың мөлшері, бірл.
Қаңтар	№1	75
	№2	12
Ақпан	№1	80
	№2	120

2.12-кесте

Жеткізілімнің орнатылған мерзімдерін бұзу динамикасы

Жеткізуші № 1			Жеткізуші № 1		
Ай	Жеткізілім-дер саны, бірл.	Кешігулер саны, тәулік	Ай	Жеткізілім-дер саны, бірл.	Кешігу-лер саны, тәулік
Қаңтар	8	30	Қаңтар	6	40
Ақпан	7	35	Ақпан	7	35

Жеткізушінің рейтингін есептеуді 2.8-кесте түрінде безендендіру керек.

Бақылау сұрақтары

1.Уақыттық қатарлардың динамикасын бағалау үшін қандай статистикалық көрсеткіштер?

2. Базистік және тізбектік өсімдер, өсу қарқыны, өсімдер қарқыны қалай анықталады?

3. Орташа өсу мен орташа өсім әдістерінің артықшылықтары мен кемшіліктерін атаңыз.

2.5. Жылжымалы орташа шаманың көмегімен уақыттық қатарларды тегістеу

Даму тенденциясын анықтау кезінде уақыттық қатарды тегістеу кең тараған тәсіл болып табылады. Тегістеу тәсілдерінің мәні уақыттық қатардың деңгейлерін тербелістері аз болатын есептеу деңгейлерімен алмастыруда. Кейде тегістеу кезінде тенденцияны анықтаудың басқа әдістерін қолдану алдында алдын ала кезең ретінде қолданады [17,18,19,22].

Жылжымалы орташалар кездейсоқ және жүйелік тербелістерді тегістеуге, үдеріс дамуындағы қолданылып жатқан тенденцияны анықтауға мүмкіндік береді, және

сондықтан уақыттық қатардың компоненттерін фильтрациялау кезінде маңызды инструмент болып табылады.

Қарапайым жылжымалы орташа шама бойынша тегістеу алгоритмін келесі қадамдар тізбегі түрінде беруге болады:

1. g тізбектелген деңгейлері кіретін қатардың g тегістеу интервалының ұзындығын анықтаймыз (g<n). Мұнда тегістеу интервалы кең болған сайын, тербелістер көп дәрежеде өзара жұтылады. Тербелістер күштірек болған сайын, тегістеу интервалы кеңірек болады.

2. Зерттеулер периодын учаскелерге бөлеміз, мұнда тегістеу интервалы 1-ге тең қадаммен қатар бойынша сырғанап тұрғандай болады.

3. Әрбір учаскені құрайтын қатардың деңгейлерінен арифметикалық орташаларды есептейміз.

4. Әрбір учаскенің ортасында тұрған қатардың нақты мәндерін орташа мәндеріне алмастырамыз.

Мұнда g тегістеу интервалының ұзындығын g=2p+1 тақ сан түрінде алу ыңғайлы, өйткені бұл жағдайда жылжымалы орташа шаманың алынған мәндері интервалдың орташа шама мүшесіне келеді.

Орташа шаманың мәндерін есептеу үшін алынған зерттеулер белсенді тегістеу учаскесінің барлық деңгейлері келесі түрде беріле алады:

$y_{t-p}, y_{t-p+1}, ..., y_{t-1}, y_t, y_{t+1}, ..., y_{t+p-1}, y_{t+p}$, ал орташа шаманың жылжымалысы келесі формуламен анықталады:

$$\widehat{y}_t = \frac{\sum_{i=t-p}^{t+p} y_i}{2p+1} = \frac{y_{t-p} + y_{t-p+1+...+y_{t+p-1}} + y_{t+p}}{2p+1} \quad (2.6)$$

мұндағы y_i – і-інші деңгейдің нақты мәні;

y_t – t мезетіндегі орташа шама жылжымалысының мәні;

2p+1 – тегістеу интервалының ұзындығы.

Тегістеу үдерісі тегістеу интервалының ұзындығы тербелістер кезеңіне тең болып алынса, уақыттық қатардағы кезеңдік тербелістерді толық жоюға алып келеді.

Мезгілдік тербелістерді жою үшін төрт және он екі мүшелік жылжымалы орташа ретінде қолданған дұрыс болады, бірақ мұнда тегістеу интервалының ұзындығының тақтылық шарты орындалмайды. Сондықтан деңгейлердің жұп саны кезінде белсенді учаскедегі алғашқы және соңғы зерттеулерін жарты салмағымен алу қабылданған:

$$\widehat{y}_t = \frac{\frac{1}{2}y_{t-p} + y_{t-p+1} + \ldots + y_{t-1} + y_t + y_{t+1} + \ldots + y_{t+p-1} + \frac{1}{2}y_{t+p}}{2p} =$$

$$= \frac{\frac{1}{2}y_{t-p} + \sum_{i=1-p+1}^{t+p-1} y_i + \frac{1}{2}y_{t+p}}{2p} . \qquad (2.7)$$

Мезгілдік тербелістерді тегістеу үшін кварталдық немесе динамикасының уақыттық қатарларымен жұмыс істеген кезде келесі жылжымалы орташаларды қолдануға болады:

$$\widehat{y}_t = \frac{\frac{1}{2}y_{t-2} + y_{t-1} + y_t + y_{t+1} + \frac{1}{2}y_{t+2}}{4} \quad , \qquad (2.8)$$

$$\widehat{y}_t = \frac{\frac{1}{2}y_{t-6} + y_{t-5} + \ldots + y_t + \ldots + y_{t+5} + \frac{1}{2}y_{t+6}}{12} . \qquad (2.9)$$

g=2p+1 белсенді учаскенің ұзындығы бар жылжымалы орташа шаманы қолданған кезде қатардың алғашқы және соңғы *p* деңгейлерін тегістеуге болмайды, олардың мәндері жоғалып кетеді.

Соңғы нүктелердің мәндерінің жоғалуы үлкен кемшілік болып табылады, өйткені зерттеуші үшін соңғы "жаңа" мәліметтердің ақпараттық құндылығы жоғары болады. Уақыттық қатардың жоғалтылған мәндерін қалпына келтіруге мүмкіндік беретін амалдардың біреуін қарастырайық. Бұл үшін келесі іс-әрекеттерді жасау қажет:

1) Соңғы белсенді учаскедегі орташа шама өсімді есептеу керек:

$$y_{t-p}, y_{t-p+1}, \ldots, y_{t-1}, y_t, y_{t+1}, \ldots, y_{t+p-1}, y_{t+p},$$

$$\overline{\Delta y} = \frac{y_{t+p} - y_{t-p}}{g-1},$$

мұндағы g – белсенді учаскенің ұзындығы;

y_{t+p} – белсенді учаскедегі соңғы деңгейдің мәні;

y_{t-p} – белсенді учаскедегі бірінші деңгейдің мәні;

$\overline{\Delta y}$ – орташа абсолюттік өсім.

2) Соңғы тегістелген мәнге орташа шаманың абсолюттік өсімін қосу арқылы уақыттық қатардың аяғында *P* тегістелген мәндерін алу.

Осындай үрдісті уақыттық қатардың бірінші деңгейлерін анықтау үшін жүзеге асыруға болады.

Қарапайым жылжымалы орташа шама әдісі динамикалық қатардың графикалық көрінісі түзуді берсе қолданылады. Тегістелетін қатардың трендінің иілісі болса, және зерттеуші үшін қысқа иілістерді сақтау тиімді болса, қарапайым жылжымалы орташа әдісін қолдану дұрыс болады.

Егер үрдіске сызықты емес даму тән болса, онда қарапайым жылжымалы орташа көптеген бұрмалауларға алып келеді. Бұл жағдайларда салмақталған жылжымалы орташаны қолданған жөн болады.

Әрбір учаскеде салмақталған жылжымалы орташа бойынша тегістеу кезінде жоғары емес реттегі полиномдар бойынша жүзеге асырылады. Көбінесе 2-інші және 3-інші реттегі полиномдар қолданылады. Қарапайым жылжымалы орташа кезінде әрбір белсенді учаскеде тегістеу түзу (бірінші реттегі полином бойынша) бойынша жасалатындықтан, қарапайым жылжымалы орташа әдісі салмақталған жылжымалы орташа әдісінің жеке жағдайы ретінде қарастырыла алады.

Қарапайым жылжымалы орташа тең салмақтары бар тегістеудің белсенді учаскесіне кіретін қатардың барлық деңгейлерін ескереді, ал салмақталған орташа әрбір деңгейге

берілген деңгейдің белсенді учаскенің ортасында тұрған деңгейден алынып тасталған салмақты қосып жазады. Салмақталған жылжымалы орташа көмегімен тегістеу келесі түрде жүзеге асырылады.

Әрбір белсенді учаске үшін $\hat{y}_t = a_0 + a_1 t + a_2 t^2 + ...,$ түріндегі полином таңдалады, оның параметрлері ең кіші квадраттар әдісі бойынша бағаланылады. Мұнда есептеу басы белсенді учаскенің ортасына жылжытылады. Мысалы, g=5 тегістеу интервалының ұзындығы үшін белсенді учаскенің деңгейлерінің индекстері келесідей болады -2, -1, 0, 1, 2. Белсенді учаскенің ортасында тұрған деңгей үшін тегістелген мән таңдалған полиномның a_0 параметрі болады.

Тегістеудің белсенді учаскесіне кіретін қатардың деңгейлері кезінде салмақтық коэффициенттерді қайта-қайта есептеу қажет емес, өйткені олар әрбір белсенді учаске үшін бірдей болады. Мұнда k-інші жұп емес дәрежелі полином бойынша салмақталған коэффициенттер (k-1) дәрежелі полином бойынша тегістеу кезіндегі коэффициенттерге тең болады. 2.13-кестеде 2-інші және 3-інші реттегі полиномдар бойынша тегістеу кезіндегі салмақтық коэффициенттер (тегістеу интервалының ұзындығына байланысты) берілген.

Салмақтар орталық деңгейге қатысты симметриялы болғандықтан, кестеде символдық жазу қолданылған: белсенді учаскенің деңгейлерінің жартысы үшін салмақ келтірілген; тегістеу учаскенің ортасында тұрған деңгейге қатысты салмақ көрсетілген. Басқа деңгейлер үшін салмақтар келтірілмеген, өйткені олар симметриялық түрде көрсетілуі мүмкін.

Мысал ретінде 5 мүшелі салмақталған жылжымалы орташа бойынша 2-інші реттегі парабола бойынша тегістеу үшін кестені қолдануды көрсетейік. Онда әрбір белсенді учаскеде орталық мән $y_{t-2}, y_{t-1}, y_t, y_{t+1}, y_{t+2}$ келесі формуламен бағаланылады:

$$\hat{y}_t = \frac{1}{35}(-3y_{t-2} + 12y_{t-1} + 17y_t + 12_{t+1} - 3y_{t-2}).$$

Келтірілген салмақтардың маңызды қасиеттерін көрсетіп кетейік:

1) Олар орталық деңгейге қатысты симметриялы.

2) Жақшадан шығарылған жалпы көбейткіші ескерілген салмақтар сомасы бірге тең.

3) Оң және теріс салмақтардың бар болуы тегістелген қисыққа трендтің қисығының әртүрлі иілмелерін сақтауға мүмкіндік береді.

Қосымша есептеулер көмегімен тегістеу интервалының g=2p+1 ұзындығы кезінде Р бастапқы және ақырғы деңгейлері үшін тегістелген мәндерді алуға мүмкіндік беретін амалдар бар.

2.13-кесте

Екінші және үшінші реттегі полиномдар бойынша тегістеу кезіндегі салмақтық коэффициенттер

Тегістеу интервалының ұзындығы	Салмақтық коэффициенттер
5	$\frac{1}{35}$ [-3, +12, +17]
7	$\frac{1}{21}$ [-2,+3,+6,+7]
9	$\frac{1}{231}$ [-21,+14,+39,+54,+59]
11	$\frac{1}{429}$ [-36,+9,+44,+69,+84,+89]
13	$\frac{1}{143}$ [-11,0,+9,+16,+21,+24,+25]

2.2-мысал. Материалдық ресурстар қажеттігін болжау. Материалдық ресурстар (МР) қажеттігін болжау үшін кіріс мәліметтер, аңғырт болжау нәтижелері және орташа тұтыну негізінде болжам жасау есептерінің нәтижелері 2.14-кестеде келтірілген [26,28,29,30,31,36].

Жылжымалы орташа бойынша болжау. Қарастырылатын мысалда МР тұтыну тербелістері жылдың бірінші жартысында екі айдан артық жалғаспады. Жылдың екінші жартысында ұзағырақ тенденциялар болды (жылдың аяғында төрт айға дейін). Қарастырылып отырған мысалдың мезгілдік тербелістер сипатын және тенденцияларын ескермей, жылжымалы орташаны есептеу интервалы ретінде екі айды таңдаймыз. Бір айдағы жұмыс күндерінің мөлшерін ескере отырып, жылжымалы орташа бойынша болжамды есептеу нәтижесі 2.14-кестеде берілген.

2.14-кесте

Материалдық ресурстар (МР) қажеттігін болжау үшін кіріс мәліметтер, қарапайым болжам нәтижелері

Ай	Бір айдағы нақты тиелім	Қара-пайым болжам	Жұмыс күндерінің саны	Бір күндегі орташа тұтыну	Тұтынудың орташа мәндері бойынша болжам	
					Бір күндегі	Бір айдағы
Қаңтар	17 246		15	1 150		
Ақпан	57 189	17 246	21	2 723	1 150	24 144
Наурыз	48 506	57 189	20	2 425	2 723	54 466
Сәуір	58 649	48 506	21	2 793	2 425	50 931
Мамыр	45 479	58 649	22	2 067	2 793	61 442
Маусым	23 835	45 479	22	1 083	2 067	45 479
Шілде	21 732	23 835	20	1 087	1 083	21 668
Тамыз	65 291	21 732	24	2 720	1 087	26 078
Қыркүйек	46 665	65 291	22	2 121	2 720	59 850
Қазан	45 345	46 665	20	2 267	2 121	42 423
Қараша	31 499	45 345	21	1 500	2 267	47 612
Желтоқсан	13 716	31 499	21	653	1 500	31 499

Салмақталған жылжымалы орташа бойынша болжау. 2.15-кестенің мәліметтері үшін алдағы кезеңнің қажеттілігін болжау үшін өткен кезеңдердің мәнділік коэффициенттерін

таңдаймыз. Соңғы кезең үшін мәнділік коэффициенті 5-ке тең деп қабылданады, соңғыдан бұрынғысы үшін – 1. Салмақталған жылжымалы орташаны есептеу 2.15-кестеде келтірілген. МР орташа тәуліктік тұтынуды (2.15-кестенің 7 бағанын қара), мысалы наурыздағы тұтынуды болжау үшін қаңтар және ақпандағы нақты орташа тәуліктік тиелім статистикасы (2.15-кестенің 4 бағаны) қажет:

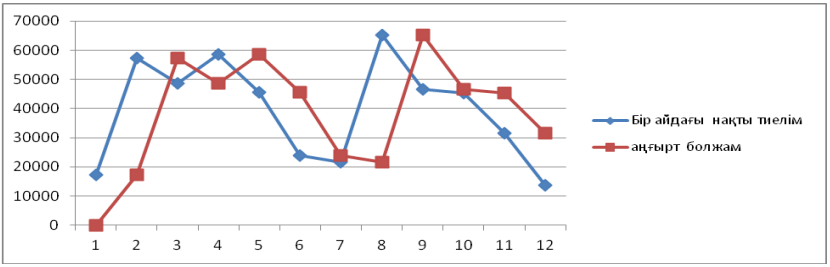
$$P = \frac{(2723*5+1150*1)}{6} \approx 2461.$$

Наурыздағы МР айлық тұтынудың болжауын алу үшін (2.15-кестенің 8 бағаны) наурыздағы орташа тәуліктік болжамды осы айдағы жұмыс күндерінің санына көбейту керек.

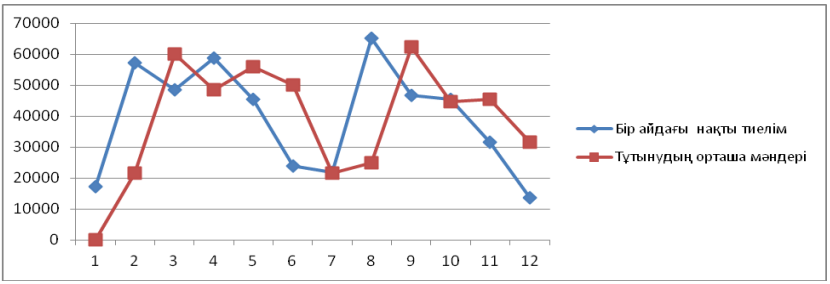
2.15-кесте

Жылжымалы орташа және салмақтанған жылжымалы орташа бойынша МР тұтынудың болжаудың мәнін есептеу

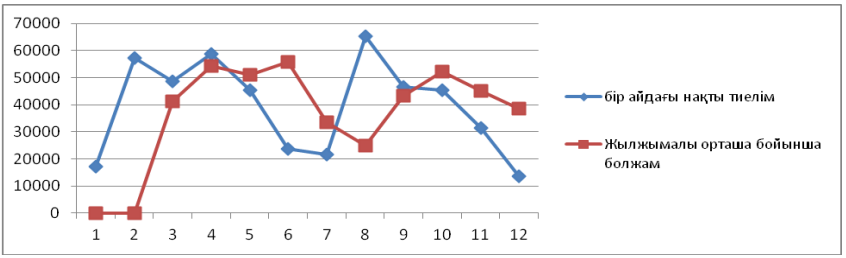
Ай	Бір айдағы нақты тиелім	Жұмыс күндерінің саны	Бір күндегі орташа тұтыну	Жылжымалы орташа бойынша болжау		Салмақтанған жылжымалы орташа бойынша болжау	
				Бір күндегі	Бір айдағы	Бір күндегі	Бір айдағы
Қаңтар	17 246	15	1 150				
Ақпан	57 189	21	2 723				
Наурыз	48 506	20	2 425	1 937	38 730	2 461	49 221
Сәуір	58 649	21	2 793	2 574	54 060	2 475	51 974
Мамыр	45 479	22	2 067	2 609	57 399	2 732	60 094
Маусым	23 835	22	1 083	2 430	53 460	2 188	48 139
Шілде	21 732	20	1 087	1 575	31 506	1 247	24 948
Тамыз	65 291	24	2 720	1 085	26 040	1 086	26 066
Қыркүйек	46 665	22	2 121	1 904	41 878	2 448	53 859
Қазан	45 345	20	2 267	2 421	48 416	2 221	44 420
Қараша	31 499	21	1 500	2 194	46 078	2 243	47 101
Желтоқсан	13 716	21	653	1 884	39 556	1 628	34 185



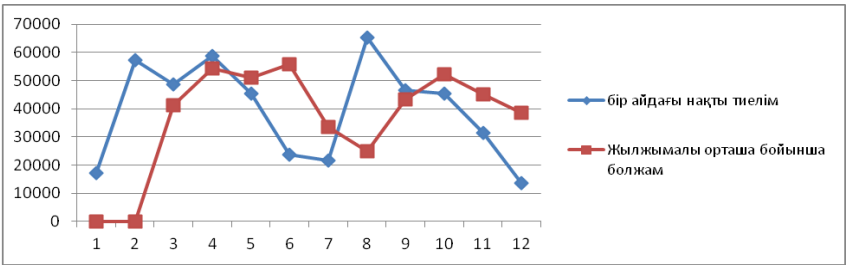
2.5-сурет. МР тұтынудың қарапайым болжау жасау нәтижелері



2.6-сурет. Тұтынудың орташа сұранысы негізінде жасалған болжау жасау нәтижелері



2.7-сурет. Жылжымалы орташа бойынша болжау жасау нәтижелері



2.8-сурет. Салмақталған жылжымалы орташа бойынша болжау жасау нәтижелері

2.3-тапсырма. А өнімі мен В өнімінің ай сайынғы сатылымдар туралы ақпарат (2.16-кесте):

2.16-кесте

Ай	2008 жыл		2009 жыл		2010 жыл	
	А өнімі	В өнімі	А өнімі	В өнімі	А өнімі	В өнімі
Қаңтар	541	256	547	325	700	298
Ақпан	458	269	625	240	458	330
Наурыз	560	120	623	256	650	320
Сәуір	670	325	523	201	658	256
Мамыр	650	245	356	295	547	356
Маусым	700	280	496	335	602	351
Шілде	630	256	630	425	500	500
Тамыз	545	345	598	395	652	520
Қыркүйек	495	355	624	325	489	340
Қазан	589	400	725	374	600	425
Қараша	660	256	562	380	457	411
Желтоқсан	630	354	540	425	511	455
Барлығы	7128	3461	6849	3976	6824	4562

Тапсырма: «Қарапайым» болжам мен ұзақ уақытты орташа әдісімен орындалған болжамның (А өнімі, тенденция жоқ) салыстырмалы сипаттамасын орындау.

Бақылау сұрақтары

1. Орташа сырғымалылар әдісі бойынша тегістеу алгоритмін келтіріңіз.
2. "Ашықтық" болжау әдісімен материалдарға қажеттілікті анықтау.
3. Ұзақ мерзімді орташа әдісімен материалдарға қажеттілікті анықтау.
4. Орташа сырғымалылар әдісімен материалдарға қажеттілікті анықтау.
5. Орта өлшеулі сырғымалылар әдісімен материалдарға қажеттілікті анықтау.

2.6. Өсу қисықтары бар модельдің көмегімен дамуды болжау

Бір өлшемді уақыттық қатарларды сипаттаудың ыңғайлы құралы уақыттың кез келген функцияларының (өсу қисықтары бар) көмегімен тегістеу болып табылады. Өсу қисығы бар динамикалық қатардың деңгейлері оның тегістелген немесе теориялық мәндерін алуға мүмкіндік береді. Бұл құбылыс динамиканың қисықпен толық сәйкес келу жағдайы бойынша орын алған деңгейлер.

Өсу қисықтарын пайдалану арқылы болжамды құрастыру үдерісіне келесі кезеңдер кіреді:

1) пішіні уақыттық қатардың өзгеру сипатына сәйкес келетін бір немесе бірнеше қисықтарды таңдау;

2) таңдалған қисықтардың параметрлерін бағалау;

3) таңдалған қисықтардың болжанатын үдеріске баламалылығын тексеру және өсу қисығын таңдау;

4) нүктелік және интервалдық болжамды есептеу.

Өсу қисықтары даму динамикасының қай типін жақсы сипаттайтындығына байланысты үш класқа бөлінеді [22].

І класқа дамудың монотонды сипаты бар және өсу шектері жоқ үрдістерді сипаттау үшін қолданылатын функциялар жатады. Бұл шарттар көптеген экономикалық көрсеткіштер

үшін дұрыс болады, мысалы, өндірістің натуралды көрсеткіштерінің көбісі үшін.

II класқа зерттелетін кезеңде өсудің шегі бар үрдісті сипаттайтын қисықтар жатады. Мұндай үрдістерді көбінесе демографияда, қызметтер мен тауарларды тұтынуды зерттегенде, ресурстарды пайдалану тиімділігін зерттеу кезінде кездестіруге болады.

Тұтынудың белгілі өнімдерін бір адамға шаққандағы тұтынуы, т.б. өсу шектері көрсетілген көрсеткіштердің мысалы бола алады.

II класқа жататын функциялар молықтыру қисықтары деп аталады. Егер молықтыру қисықтарының иілім нүктелері болса, онда өсу қисықтарының III класты S-пішінді қисықтарға жатады.

Бұл қисықтар екі бірінен кейін бірі жүретін көшкін пішінді үрдісті сипаттайды: біреуі дамудың жеделдетуімен, екіншісі баяулатумен.

S-пішінді қисықтар демографиялық зерттеулерде, сақтандыру есептеулерінде, ғылыми-техникалық дағдарысты болжау есептерін шешу кезінде, өнімнің жаңа түріне сұранысты анықтау кезінде қолданылады.

Қисықты таңдау мәселесі қатарды тегістеу үшін өзекті болады.

Бұл есепті шешудің бірнеше тәсілдемесі болады, бірақ олардың барлығы қолданылатын өсу қисықтарының негізгі қасиеттерімен танысуды қажет етеді. Сондықтан іс-жүзінде жиі қолданылатын қисықтардың арнайы кластық сипаттамаларына тоқтайық. I кластағы өсу қисықтары арасында полиномдар класын көрсету керек

$$y_t = a_0 + a_1 t + a_2 t^2 + ... + a_p t^p, \qquad (2.10)$$

мұндағы a_i $(i=0,1, ... ,p)$ – көп мүшенің параметрлері,

t – тәуелсіз айнымалы (уақыт).

Жоғары емес дәрежелі полиномдардың коэффициенттері динамикалық қатардың мазмұнына байланысты нақты интерпретациясы болуы мүмкін. Мысалы, оларды (a_1) өсу жылдамдығы, (a_2) өсудің жеделдеуі, (a_3) жеделдетудің өзгерісі, t

=0 кезінде (a_0) қатардың бастапқы деңгейі. Көбінесе экономикалық есептерде үшінші реттен жоғары емес полиномдар қолданылады. Трендті анықтау үшін жоғары дәрежелі полиномдарды қолдану дұрыс емес, өйткені осылай алынған аппроксималдаушы функциялар кездейсоқ ауытқуларды көрсетеді (бұл тенденцияның мағынасына қайшы келеді).

$y_t=a_0+a_1t$ бірінші дәрежелі полином графикте түзумен көрсетіледі және уақытта тепе-тең дамитын үрдісті сипаттау үшін қолданылады.

$y_t=a_0+a_1t+a_2t^2$ екінші дәрежелі полином үрдісі тең жеделденетін болған жағдайда қолданылады.

Негізінен, $a_2>0$ параметрі болса, парабола бұтақтары жоғары, егер $a_2<0$ болса, онда төмен бағытталған. a_0 және a_1 параметрлері парабола пішініне әсер етпейді, ол оның тек орналасуын анықтайды.

Үшінші дәрежелі полиномның түрі $y_t=a_0+a_1t+a_2t^2+a_3t^3$.

Бұл полиномның ордината өсімінің белгісі бір немесе екі рет өзгере алады.

Полиномдардың ерекшелігі – өсімдердің (y_t) ордината мәндерінен тәуелділігінің жоқтығы. (2.10) моделінде параметрлерді бағалау ең кіші квадраттар әдісімен анықталады. Оның мәні деңгейлердің есептелген мәндерінің нақты мәндерден ауытқуларының квадраттар сомасы минималды болатындай параметрлерді «іздеуде». Осылайша, бұл бағалар келесі өрнекті минимизациялау нәтижесінде анықталады:

$$\sum_{t=1}^{n}(y_t-\hat{y}_t)^2\to\min, \qquad (2.11)$$

мұндағы y_t – уақыттық қатардың нақты мәні;

$\hat{y}_t$ – есептелген мәні;

n – уақыттық қатардың ұзындығы.

(2.11) өрнекті минимизациялау нәтижесінде алынған теңдеулер жүйесін келтіреміз:

$$\sum_{t=1}^{n} y_t = a_0 \times n + a_1 \sum t + a_2 \sum t^2 + ... + a_p \sum t^p$$

$$\sum y_t \times t = a_0 \sum t + a_1 \sum t^2 + a_2 \sum t^3 + ... + a_p \sum t^{p+1}$$

... .

$$\sum y_t \times t^{p-1} = a_0 \sum t^{p-1} + a_1 \sum t^p + a_2 \sum t^{p=1} + ... + a_p \sum t^{2p-1}$$

$$\sum y_t \times t^p = a_0 \sum t^p + a_1 \sum t^{p+1} + a_2 \sum t^{p+2} + ... + a_t \sum t^{2p}$$

(2.12) жүйесі $a_0, a_1, ..., a_p$ белгісіз шамалары ретінде $(p+1)$ коэффициенттерден тұратын $(p+1)$ теңдеулерден тұрады. Бұл жүйенің шешімі ізделіп отырған коэффициенттердің бағаларын есептеуге мүмкіндік береді.

Жоғары емес дәрежелі полиномдарды бағалау үшін жүйелер қарапайым болып келеді. Мысалы, түзудің параметрлерін бағалау үшін нормальді теңдеулер келесі түрде беріледі:

$$\begin{aligned} \sum y_t &= a_0 n + a_1 \sum t; \\ \sum y_t \times t &= a_0 \sum t + a_1 \sum t^2. \end{aligned} \qquad (2.13)$$

Ізделіп отырған параметрлеріне қатысты бұл жүйенің шешімі келесі өрнекті береді:

$$a_1 = \frac{\sum y_t \cdot t - \frac{\sum y_t}{n} \cdot \sum t}{\sum t^2 - \frac{(\sum t)^2}{n}}; \qquad a_0 = \frac{\sum y_t}{n} - a_1 \frac{\sum t}{n}.$$

2-інші реттегі параболаны есептеу үшін осындай нормальді теңдеулер жүйесін келтіреміз:

Түзудің теңдеуі:

$$a_0 = \frac{\sum y_t}{n}; \qquad a_1 = \frac{\sum y_t \cdot t}{\sum t^2}. \qquad (2.15)$$

Параболаның теңдеуі:

$$a_0 = \frac{\sum y_t}{n} - \frac{\sum t^2}{n}\left|\frac{n\sum y_t \cdot t^2 - \sum t^2 \sum y_t}{n\sum t^4 - \left(\sum t^2\right)^2}\right| \quad ;$$

$$a_1 = \frac{\sum y_t \cdot t}{\sum t^2}; \qquad a_2 = \frac{n\sum y_t \cdot t^2 - \sum t^2 \cdot \sum y_t}{n\sum t^4 - \left(\sum t^2\right)^2}.$$

Полиномдарға қарағанда, экспоненциалды қисықтар үшін өсімнің функцияның өзінің шамасынан тәуелділігі тән болады. Бұл қисықтар өсім функцияның жеткен деңгейіне байланысты болатын "көшкін тәрізді" сипаттағы үрдістеді жақсы сипаттайды.

Қарапайым экспоненциалды (көрсеткіштік) қисықтық түрі келесідей болады:

$$y_t = a \cdot b^t. \qquad (2.17)$$

Егер b>1 болса, онда қисық t өсуімен өседі, егер b<1 болса, түседі.

a параметрі дамудың бастапқы шарттарын сипаттайды, және *b* параметрі – өсудің тұрақты қарқынын сипаттайды.

Шындығында, өсудің қарқыны: $T_t = \frac{y_t}{y_{t-1}} 100° / \circ.$

Берілген жағдайда $T_t = \frac{a \cdot b^t}{a \cdot b^{t-1}} 100° / \circ = b \cdot 100° / \circ = const.$

Осыған сәйкес өсу қарқындары да тұрақты болады.

$K_t=T_t-100\%=const.$ Бұл функцияның ординатасының логарифмі t-ға сызықты тәуелді, ол үшін (2.17) өрнегін логарифмдейміз

$$\log y_t = \log a + t \log b.$$

Айталық $\log a=A$; $\log b=B$. Онда $\log y_t = A+tB$.

Енді белгісіз параметрлерді бағалау үшін (2.13) түзуі үшін нормальді теңдеулер жүйесін қолдануға болады.

Басқаша айтқанда, минимизациядан нормальді теңдеулер құрылады:

$$\sum \left(\log y_t - \log \hat{y}_t\right)^2 \to \min.$$

Осыған сәйкес нормальді теңдеулерде нақты деңгейлердің орнына олардың логарифмдері қолданылады:

$$\sum \log y_t = n \cdot A + B \sum t;$$
$$\sum (\log y_t \cdot t) = A \sum t + B \sum t^2. \qquad (2.18)$$

Белгісіз А және В параметрлерін табамыз. *A =log a* және *B =log b* мәндерін біле отырып, *a* және *b* мәндерін табамыз, және потенциалдау көмегімен қатарды тегістеуге қызмет ететін көрсеткіштік функцияны аламыз.

Белгісіз параметрлерді бағалауға қолданылатын мұндай тәсілдеме өзінің әмбебаптылығымен көзге түседі. Бірақ параметрлердің алынған бағалары ығыстырылған болып шығатынын ескеру керек, өйткені есептеуге кіріс мәліметтер емес олардың логарифмдері қатысады. Динамикалық қатардың көршілес деңгейлері арасындағы айырма үлкен болған сайын, ығыстыру үлкен болады. Осығын ұқсас жағдайларда ығыстыруға сызықтық емес ең кіші квадраттар әдісі алып келмейді.

Экспоненциалды қисықтың күрделірек варианты логарифмдік парабола болады

$$y_t = a \times b^t \times c^{t^2}. \qquad (2.19)$$

(2.19.) өрнегін логарифмдеп, келесі параболаны аламыз:

$$\log y_t = \log a + t \log b + t_2 \log c,$$

осылайша, логарифмдік параболаның параметрлерінің бағасын (2.14) параболасы үшін нормальді теңдеулер жүйесін қолдана отырып, ең кіші квадраттар әдісі көмегімен жүзеге асыруға болады. Мұнда алынған бағалардың ығыстыруы туралы жоғарыда жасалынған барлық ескертулер орынды болады [27,30,31].

2.3-мысал. 2005–2010 жж. аралығында жергілікті қоймадан жүктерді тасымалдау көлемінің динамикалық қатары белгілі (2.17-кесте). 2013 жылға арналған тасымалдаудың болжауын жасау керек.

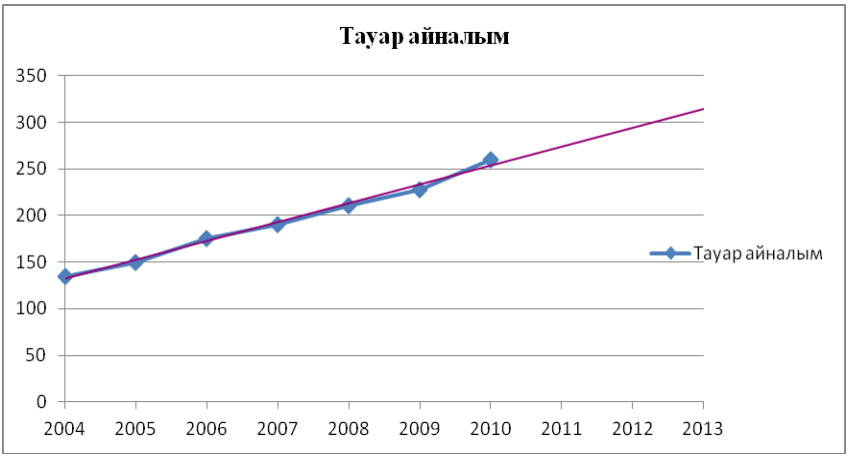
Келесі формулалармен *a, b* параметрлерін анықтаймыз:

$$b = \frac{\sum xy}{\sum x^2}; \quad a = \frac{\sum y}{n}.$$

2.17-кесте

2005–2010 жж. аралығында тасымалдау көлемі (мың, тг)

Жылдар	2004	2005	2006	2007	2008	2009	2010
Тауар айналым	135	150	175	190	211	228	260



2.10-сурет. 2005–2010 жж. кезеңіндегі тасымалдау көлемінің өзгеру динамикасы: 1 – нақты мәліметтер; 2 – есептелген мәліметтер

2.18-кесте

Жергілікті қойманың тасымалдау көлемін тегістеу және болжау үшін парабола теңдеуінің параметрлерін есептеу

Жылдар, x	y	X	xy	x^2	y^=192+20,8x
1	2	3	4	5	6
2004	130	-3	-390	9	129,6
2005	150	-2	-300	4	150,4
2006	175	-1	-175	1	171,2
2007	190	0	0	0	192,0
2008	211	1	211	1	212,8

2009	228	2	456	4	233,6
2010	260	3	780	9	254,4
	1344	0	582	28	1344
2011		4			275,1
2012		5			295,9
2013		6			316,7

Әрбір жыл үшін қатардың теориялық деңгейін есептейміз (2.18-кесте, 6 баған). Әрбір жыл бойынша салыстырылған 2 және 6 графалары есептеу деңгейлерінің нақты деңгейлерден кішігірім ауытқуларын көрсетеді, бұл математикалық теңдеуді таңдаудың дұрыстығын көсетеді.

Тауар айналымын болжау үшін көрсетілген сандардың артынан жүретін 3 графаны жалғастыру керек. Біздің жағдайымызда бұл 3, ары қарай қарастырылған кезең 4, 5, 6 болады және т.с.с. 2013 жылы x=6, онда $y_{2013}=192+20,8*6=316,7$.

2.4-тапсырма. 2008, 2009 жж. аймақтық қойманың тауар айналымының болжауын жасаңыз. Осы қойманың 2009 ж. материал ағындарының болжауын жасаңыз.

2.19-кесте

Тауар айналымды есептеу үшін кіріс мәліметтер

Көрсеткіш*	Өлшем бірлігі	2002	2003	2004	2005	2006	2007
Т	млн тг.	80	120	145	185	210	240
Q	мың. т.	245	340	600	840	616	1850
Нр	м/млн.тг	2400	3000	4500	4800	4500	5700
М	%	32	27	19	16	11	9
У	%	90	92	85	88	86	87

*Т – тауарайналым; Q – тасымалдау көлемі; Нр – тауар айналымның 1 млн теңгесіне жатқызылған тасымалдау көлемінің үлестік көрсеткіші; М – автокөлікпен жүкті орталықтандырылмаған тасымалдаулардың үлестік салмағы; У – жүкті тиеу және түсіру кезіндегі жұмыстарды механизациялау деңгейі.

2.5-тапсырма. 2010–2016 жж. аралығында жергілікті қоймадан жүктерді тасымалдау көлемінің динамикалық қатары белгілі (2.20-кесте). 2017 жылға арналған тасымалдаудың болжауын жасау керек [37,38].

2.20-кесте

2010–2016 жж. аралығында тасымалдау көлемі (млн.долл.)

Жылдар	2010	2011	2012	2013	2014	2015	2016
Сыртқы сауда айналымы	91 397,5	121 241,7	132 807,2	133 506,0	120 755,3	76 523,5	61 950,4

Бақылау сұрақтары

1. Өсім қисығы көмегімен болжам қалай жасалады?
2. Өсім қисықтары қандай түрлерге бөлінеді.
3. Полином коэффициенттері қандай әдістермен анықталады?

2.7. Логистикалық үрдістерді болжауда өсу қисықтарының модельдерін қолдану

Қисықтардың барлық қарастырылған типтері "қанағаттандыру" жоқ монотонды өспелі және кемімелі үдерістерді сипаттау үшін қолданылады [22].

Үрдісті деріс "қанағаттандырумен" сипатталса, оны нөлден өзгеше асимптотасы бар қисық көмегімен сипаттау керек. Мұндай қисықтың мысалы ретінде модифицияланған экспонента арқылы көрсетуге болады:

$$y_t = k + ab^t , \qquad (2.20)$$

мұнда y = k горизонтальді асимптота болып табылады.

Егер a параметрі теріс болса, онда асимптота қисықтан жоғары орналасады, егер a оң болса, төмен орналасады.

Экономикалық есептерді шешу кезінде a<0, b<1 болатын қисықпен жұмыс істеу керек болады [18–20].

Экономикалық есептерді шешу кезінде асимптотаның мәнін болжанатын үрдістің қасиеттеріне сүйене отырып анықтауға болады (мысалы, жабдықты қолдану коэффициенті 1-ден аспау керек). Кейде асимптотаның мәні сараптау жолымен беріледі. Осындай жағдайларда қисықтың параметрлері теңдеулерді сызықтық түрге келтіргеннен кейін ең кіші квадраттар әдісі көмегімен анықтауға болады:

$$y_t - k' = a \cdot b^t, \qquad (2.21)$$

мұндағы k – асимптотаның берілген мәні.

(2.21) өрнегін логарифмдейік:

$$\log(y_t - k') = \log a + t \log b^t.$$

Енді (2.18) нормальді теңдеулер жүйесін қолданып, log *a* және log *b* параметрлерін анықтауға болады.

Модифицияланған экспонентаның параметрлерін бағалау үшін сызықтық емес ең кіші квадраттар әдісін де, сондай-ақ есептеуі жеңіл, ал бағаларының тиімділігі азырақ болатын басқа әдістердің бірқатарын қолдануға болады.

Осылайша, модифицияланған экспонента дамуына шектелген фактор әсер ететін үрдісті жақсы сипаттайды, мұнда бұл әсердің өсуі жеткен деңгейдің өзгеруімен бірге өседі.

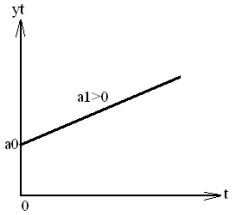
Егер шектелген фактордың әсері белгілі мезеттен (иілім нүктесінен) кейін көріне бастаса, оған дейін ол әйтеуір бір экспоненциалды заңдылықпен дамыса, онда тегістеу үшін S-пішінді қисықтар қолданылады.

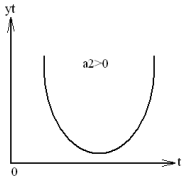
Гомперц қисығы және логистикалық қисық немесе Перл-Рид қисығы олардың ішіндегі ең белгілілері болып табылады.

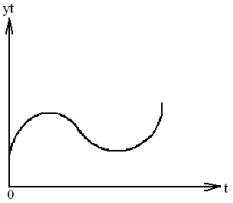
Гомперц қисығының түрі: $y_t = k \times a^{b^t}$.

Қисық симметриялы емес. Егер log *a* <0, онда қисықтың түрі S-пішінді болады, мұнда *k-ға тең* асимптота қисықтан жоғары өтеді. Егер log *a* >0, *k-ға тең* асимптота қисықтан төмен өтеді, ал қисықтың өзі монотонды өзгереді: *b*<1 кезінде – монотонды кемиді; *b*>1 кезінде – монотонды өседі.

Экономикалық есептерді шешу кезінде log a <0 және b<1 (2.11-сурет) болатын қисықтың варианты қызық болады.



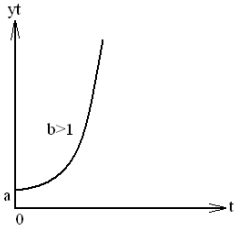


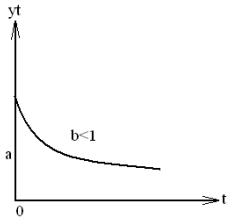


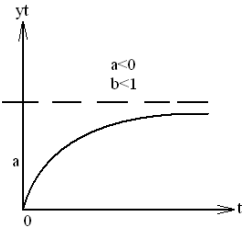
1) бірінші реттегі полином $y_t=a_0+a_1t$;

2)екінші реттегі полином $y_t=a_0+a_1t+a_2t^2$;

3) үшінші реттегі полином $y_t=a_0+a_1t+a_2t^2+a_3t^3$;

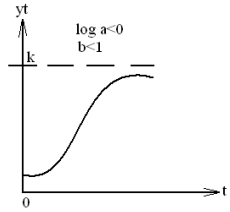


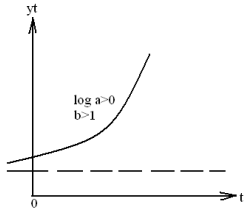


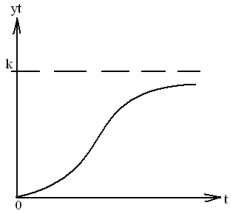


4) экспонента $y_t=a\times b^t$;

5) модифицияланған экспонента $y_t=k+a\times b^t$;







6) Гомперц қисығы

7) логистикалық қисық кривая $y_t=k\times a^{bt}$

$$\frac{1}{y_t} = k + a \times b^t$$

2.11-сурет. Логистикалық үрдістерді болжау үшін пайдалынатын өсу қисықтарының түрлері

Логистикалық қисықтың теңдеуі модифицияланған экспонентада y_t кері $\frac{1}{y_t}$ шамамен алмастыру арқылы алынады. *Қисықтың* теңдеуі $\frac{1}{y_t} = k + ab^t$ шамамен алмастыру арқылы алынады.

Логистикалық қисық теңдеуін жазудың басқа да түрі де қолданылады: $y_t = \frac{k}{1+be^{-at}}$.

$t \to -\infty$ кезінде ордината нөлге, ал $t \to \infty$ - кезінде k параметрінің мәніне тең асимптотаға ұмтылады. Қисық келесі $t = ln\ b:\ a;\ y_t = k : 2$ координаталары бар иілім нүктесіне қатысты симметриялы.

Графиктен көрініп тұрғандай, логистикалық функция алдымен жедел қарқынмен өседі, сосын өсу қарқыны баяулайды, соңында өсу толығымен тоқтайды.

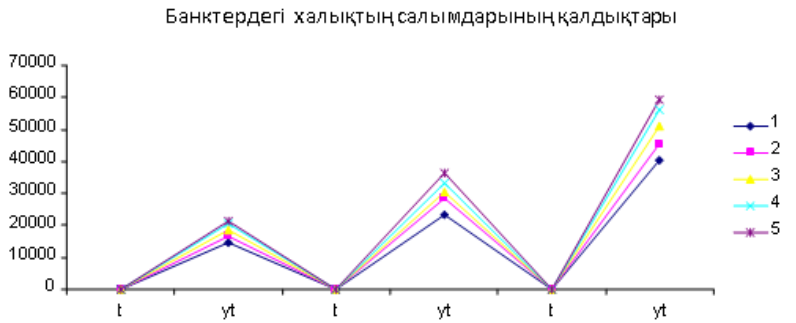
Осылайша, біз экономикалық зерттеулерде жиі қолданылатын өсу қисықтарының түрлерін қарастырдық. Бұл қисықтардың анықталған ерекшеліктері мен қасиеттері қисықтардың типін таңдау есептерін шешу кезінде көмектеседі.

2.4-мысал. 2.21-кестеде 15 айдағы банктердегі халықтың салымдарының қалдықтары туралы мәліметтер берілген. Салымдар қалдықтарының әрбір айдың басындағы мөлшері көрсетілген (2.12-сурет).

2.21-кесте

Банктердегі халықтың салымдарының қалдықтары (млн. тг.)

T	y_t	t	y_t	t	y_t
1	14717	6	23342	11	40524
2	16642	7	28317	12	45416
3	18504	8	30624	13	50857
4	20376	9	33408	14	56024
5	21321	10	36505	15	59381



2.12-сурет. Банктердегі халықтың салымдарының қалдықтары

Қатардың тенденциясы келесі түрде сипатталады деп ұйғарып, 16-шы айдың басына банктердегі халықтың салымдарының қалдықтарының болжауын есептеу:

а) сызықтық модель $y_t = a_0 + a_1 t$;

б) параболалық модель $y_t = a_0 + a_1 t + a_2 t^2$;

в) көрсеткіштік модель $y_t = ab^t$.

Шешімі

а) Сызықтық тренд коэффициенттерін есептеу үшін (2.13) қатардың ортасына координаттар басын ауыстырғаннан кейін нормальді теңдеулер жүйесінен алынған өрнектерді қолданамыз.

2.22-кестеде қажетті қосымша есептеулер келтірілген және тренд параметрлері графигі құрылған. (2.13-сурет)

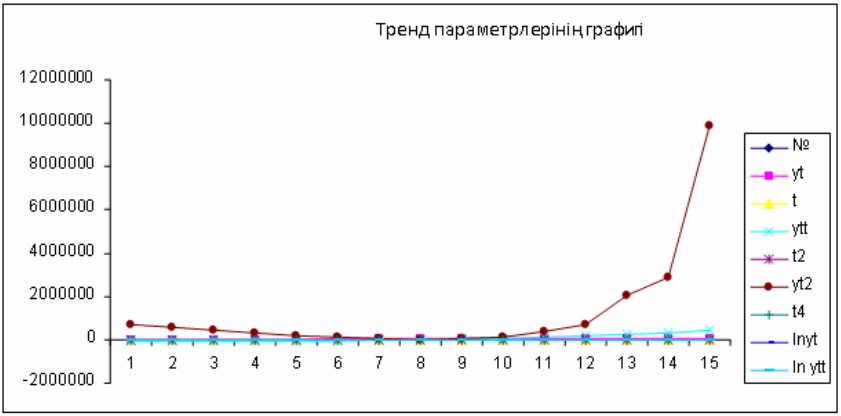
2.22-кесте

Трендтер параметрлерін есептеу

№	y_t	t	$y_t t$	t^2	y_t^2	t^4	$\ln y_t$	$\ln y_t t$
1	14717	-7	-103019	49	721133	2401	9,596759	-67,1773
2	16642	-6	-99852	36	599112	1296	9,719685	-58,3181
3	18504	-5	-92520	25	462112	625	9,825742	-49,1287
4	20376	-4	-81504	16	326016	256	9,922113	-39,6885
5	21321	-3	-63963	9	191889	81	9,967448	-29,9023
6	23342	-2	-46684	4	93368	16	10,05801	-20,116

2.22-кестенің жалғасы

7	28317	-1	-28317	1	28317	1	10,25122	-10,1512
8	30624	0	0	0	0	0	10,32954	0
9	33408	1	33408	1	33408	1	10,41655	10,41655
10	36505	2	73010	4	146020	16	10,5052	21,01041
11	40524	3	121572	9	364716	81	10,60965	31,82895
12	45416	4	181664	16	726656	256	10,72362	42,89448
13	50857	5	254285	25	1271425	625	10,83677	54,18387
14	56024	6	336144	36	2016864	1296	10,93354	65,60121
15	59381	7	415667	49	2909669	2401	10,99173	76,94211
	495958		899891	280	9891193	9352	154,6876	28,29541



2.13-сурет. Тренд параметрлерінің графигі

(2.22) кестеге сәйкес келесі есептеулер жүргіземіз:

$$a_0 = \frac{495958}{15} = 33063{,}66$$

$$a_1 = \frac{899891}{280} = 3213{,}896$$

Бұдан шығатыны, сызықтық тренд теңдеуі:

$$y_t = 33063{,}866 + 3213{,}896 \cdot t.$$

Осы модельге сәйкес t=0 қатардың орташа деңгейінің бағасы 33063,9 млрд. тг, ал халықтың салымдарының қалдықтарының орташа айлық өсімі 3213,9 млрд. тг құрады.

Алынған модель негізінде бір нүкте болашаққа болжау жасау үшін оған уақыттық қатардың тиісті мәнін қою керек, яғни t=8. (Егер модельдің коэффициенттерінің бағалары координаттар басын сол қатардың ортасына ауыстырусыз алынған болса, онда уақыттық қатардың t=16 параметрінің мәнін қою керек еді).

Болжау келесідей болады:

$$\hat{y} = 33063{,}866 + 3213{,}896 \times 8$$

$$\hat{y} = 58775 \text{ млн., тг}$$

б) параболалық тренд коэффициенттерін есептеу үшін координаттар басын қатардың ортасына ауыстырғаннан кейін нормальді теңдеулер жүйесінен алынған өрнектерді қолданамыз (2.15).

Аралық есептеулер 2.22-кестеде көрсетілген.

$$a_1 = \frac{899891}{280} = 3213{,}896$$

$$a_2 = \frac{15 \times 9891193 - 280 \times 495958}{15 \times 9352 - (280)^2} = 153{,}517$$

$$a_0 = 33063{,}866 - \frac{280}{15} \times 153{,}517 = 30198{,}16$$

Бұдан шығатыны, параболалық тренд теңдеуінің түрі келесідей болады:

$$\hat{y}_t = 30198{,}16 + 3213{,}896t + 153{,}517t^2.$$

Көрсеткіштің болжауын анықтау үшін алынған модельге уақыттық қатардың тиісті мәнін қою керек (t=8).

Болжау келесідей болады:

$$\hat{y}_t = 30198{,}16 + 3213{,}896 \times 8 + 153{,}517 \times 8^2$$
$$\hat{y} = 65734 \text{ млн., тг}$$

в) Көрсеткіштік функциямен сипатталатын тренд параметрлерін анықтау үшін (2.18), (2.14) өрнектерін қолданамыз:

$$\ln a = \frac{154{,}6876}{15} = 10{,}3125$$
$$\ln b = \frac{28{,}29541}{280} = 0{,}1011$$

Потенциалдауды жүргізіп, келесі мәндерді аламыз: a=30106,61; b=1,11.

Бұдан шығатыны, тренд теңдеуінің түрі келесідей болады:

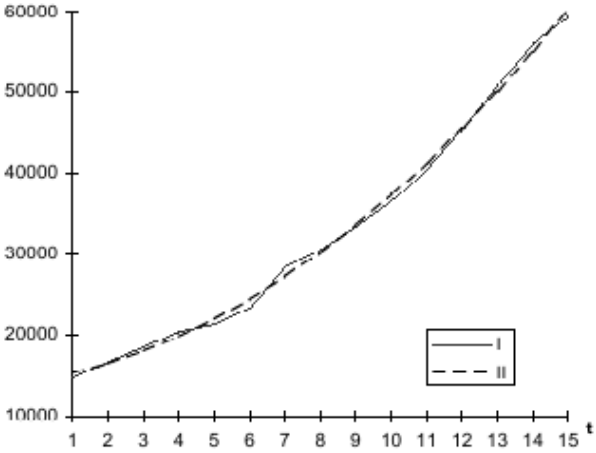
$$\hat{y}_t = 30106{,}61 \times 1{,}11^t.$$

Осы модельге сәйкес халықтың салымдарының қалдықтарының орташа айлық өсу қарқыны 111% құрады. Есептеу басы деп қабылданған нүктеде (t=0), тренд мәні 30106,61 млн. тг құрады. Бір айдан кейінгі халық салымдары қалдықтарының болжауын анықтау үшін алынған модельге t=8 мәнін қоямыз:

$$\hat{y}_t = 30198{,}16 \times 1{,}11^8$$
$$\hat{y} = 69382 \text{ млн., тг}$$

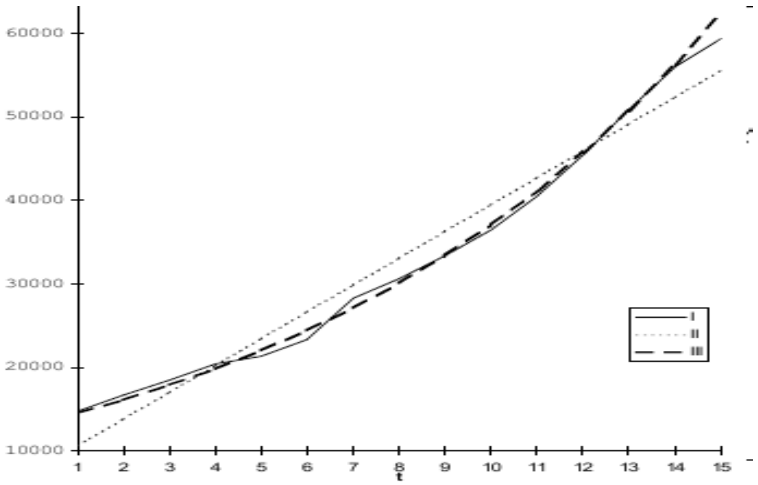
2.13-суретте сызықтық, параболалық, көрсеткіштік үш трендтік модельдері негізінде алынған деңгейлердің тегістелген мәндері және кіріс уақыттық қатардың графигі көрсетілген. Графикалық талдау сызықтық модельді адекватты деп айтуға болмайтынын көрсетті. Көрсеткіштік функция бойынша есептелген модель де шындықтан алыс. Бәрінен де нақты мәндерге параболалық модель бойынша тегістелген деңгейлер жақын, бірақ мұнда болжау мәндері жоғарылатылған.

y(t)-млн.тг



2.13, а-сурет. Уақыттық деңгейлердің нақты (I) және парабола бойынша (II) тегістелген мәндері

y(t)-млн.тг



2.13, б-сурет. Уақыттық деңгейлердің нақты (I) және түзу (II) және III – көрсеткіштік функция) бойынша тегістелген мәндері

Өсудің қисықтарын таңдау тәсілдері.

Өсудің қисығының пішінін таңдау үрдісін жеңілдететін бірнеше тәжірибелік тәсілдемелер бар.

Олардың ең қарапайым жолы – уақыттық қатардың графикалық көрінісіне сүйенетін бейнелеу тәсілі. Үрдістің нақты дамуына сәйкес келетін өсу қисығы таңдалынып алынады. Егер кіріс қатардың графигінде даму тенденциясы жеткіліксіз анық көрінсе, қатардың кейбіріне стандартты түрлендірулер жасауға болады, сосын түрлендірілген қатардың графигіне жауап беретін функцияны таңдау керек. Статистикалық өңдеудің қазіргі пакеттерінде әртүрлі масштабтарда мәліметтердің стандартты түрлендірулердің үлкен арсеналы және графикалық көрініс үшін көп мүмкіндіктер бар. Осының бәрі зерттеуші үшін берілген кезеңді жеңілдетуге мүмкіндік береді.

Статистикалық әдебиетте параболалық типтегі қисықтарды таңдау кезінде көмектесетін тізбектелген айырмалар әдісі сипатталған. Бұл әдіс келесі ұйғарымдар жасау кезінде қолданыла алады: уақыттық қатардың деңгейлерінің дисперсиясы тұрақты, математикалық күтімі о-ге тең үлестірілудің нормальді заңына бағынатын жүйелік құрама және кездейсоқ компонентаның сомасы ретінде беріле алады:

$$\Delta y_t = y_t - y_{t-1}$$

$$\Delta^2 y_1 = \Delta y_{t-1}$$

және т.с.с.

Есептеу айырмалар жуық шамада тең болғанша жүргізіле береді. Айырманың реті тегістелетін полиномның дәрежесі ретінде алынады. Өсу қисықтарын таңдау кезінде өсім сипаттамаларының әдісі көмектесе алады.

Бұл әдісті қолдану арқылы қисықты таңдау үрдісті келесі қадамдардан тұрады:

1) жылжымалы орташа бойынша қатарды тегістеу;

2) орташа өсуді анықтау;

3) өсудің туынды сипаттамаларын есептеу.

Қисықтардың көптеген түрлері үшін t-ға қатысты сызықты өзгеретін және тұрақты болатын өсімдердің түрлендірулері табылған. Осыған байланысты өсу сипаттамаларының

қатарларын зерттеу кіріс уақыттық қатарлардың даму заңдарын анықтау кезінде көмекке келеді.

Берілген әдіс тізбектелген айырма әдісімен салыстырғанда әмбебап болып келеді [22–25].

Бірақ іс жүзінде қисық пішінін таңдауға тегістеумен алынған деңгейдің нақты мәндерінің есептелген мәндерінен ауытқуларының квадратының сомасы ретінде алынған критерий мәндеріне сүйене отырып жақындайды. Қарастырылған қисықтардың ішінен критерийдің минималды мәні сәйкес келетін түрі таңдалады, өйткені кретерий мәні кіші болған сайын, зерттеулердің мәліметтері қисыққа соғұрлым жақын орналасады.

Бұл тәсілді қолданғанда, бірнеше мезеттерге мән беру керек. Біріншіден, m нүктелерден тұратын қатарға барлық m нүктелерден өтетін (m-1) дәрежелі көпмүшені таңдауға болады. Одан басқа, барлық осы нүктелерден тұратын жоғары дәрежелі көпмүшелер жиыны бар. Бұл көпмүшелер үшін критерий мәні 0-ге тең болады, бірақ мұндай қисық тенденцияға, болжау мақсаттарына да онша жарамды болмайтыны анық. Осылайша, бұл тәсілдің қолданылуы екі кезеңмен өтуі керек.

Бірінші кезеңде есептің мазмұнды талдауына сүйене отырып, қолайлы функциялар шектеледі.

Екінші кезеңде критерий мәндері есептеледі және оның негізінде қажетті өсу қисығы таңдалады. Зерттелетін үрдістің мазмұнды талдау қажеттігі келесі мысалдармен көрсетіледі.

Болашақтың динамика қатары экспоненциалды қисық көмегімен жақсы сипатталған делік. Бірақ, логистикалық қисықтың бірінші жартысы да экспонентамен берілген. Сондықтан келешекте экспоненциалды тенденция туралы гипотезаны мазмұнды талдау өткізілгеннен кейін ғана қолдануға болады, оның барысында шарттардың берілген жиыны кезінде “толықтыру” болуы мүмкін бе деген сұрақтарға жауап беру керек. Мысалы, өндіріс үрдісі материалдық ресурстармен немесе өндірістік қуаттылықтармен шектелуі мүмкін.

Берілген критерий бойынша түзу ең жақсы функция болатын жағдай болуы мүмкін, бірақ оның негізінде алынған болжаудың мән теріс болады. Егер көрсеткіштің экономикалық

мағынасынан теріс мәндер мүмкін еместігі шығатын болса (мысалы, шығарылған өнім көлемін болжау кезінде), онда әрине, берілген критерий "сәтті" азырақ, бірақ көрсеткіштің мазмұнды мағынасына сәйкес келетінін таңдап, бұл функциядан бас тарту керек.

Мәліметтерді статистикалық өңдеу және уақыттық қатарларды талдаудың қазіргі пакеттерінде өсу қисықтарының кең де көп түрлері көрсетілген және олар оқу үрдісінде көп қолданылады, 16-сутеттегі өсу қисықтары жүзеге асырылған. Мұнда қолданушы үшін ыңғайлы жұмыстың бірнеше тәртіптері болуы мүмкін. Бұл қисықтардың ішінен арнайы функцияны таңдауға болады, және параметрлердің бағалары, қалдықтардың сипаттамалары, интервалды және нүктелік болжаулар кіретіндігіне толық талдау алуға болады. Экранда бірнеше функцияларды көрсетуге болады, онда талдауда барлық тапсырылған параметрлердің бағалары және олардың әрқайсысы үшін критерийдің мәні болады. Критерий ретінде орташа квадраттық қате таңдалады:

$$S = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)}{n}},$$

мұндағы y_t – қатардың нақты мәні;

$\hat{y}_t$ – қатардың тегістелген мәні;

n – қатардың ұзындығы.

Ақырында өсу қисықтарын таңдауға «қатаң» ұсыныстар жоқ екенін айтамыз. Әсіресе келешекте экстраполяризациялау үшін анықталған заңдылықтар үшін алынған функцияны қолдану кезінде абайлап қарау керек. Өсу қисықтарын пайдалану болжанатын кезеңде анықталған тенденцияны сақтау туралы ұйғарымға негізделу керек.

Берілген бөлімде қарастырылған статистикалық әдістер мен тәсілдер өсудің тиісті қисығын таңдауды жүзеге асыру кезінде зерттеушіге көмектесе алады [21,23,25].

2.5-мысал. Көтерме сауданың тауар айналымын болжау моделі.

Жұмысты орындау реті:

1. 2.23-кестеде келтірілген кіріс мәліметтері негізінде өзгерісі

y (Tx) = a + b/x гипербола теңдеуімен берілген 2001–2005 жж. нақты мәндердің динамикалық қатарын тегістеу көмегімен тауар айналымның

(2006) болашағы бар көрсеткішін анықтау керек.

2. Көтерме ұйымның нақты және теориялық тауар айналым динамикасының графиктерін тұрғызу керек.

2.23-кесте

	Көрсеткіштер /жыл	D
У осі	Тауар айналым / 0,1 млн. тг	93
	Тауар айналым / 0,2 млн. тг	114
	Тауар айналым / 0,3 млн. тг	163
	Тауар айналым / 0,4 млн. тг	214
	Тауар айналым / 0,5 млн. тг	263
	Сомасы	847

Жұмысты орындау әдістемесі:

1. *a* және b тауар айналым теңдеулерінің параметрлерін анықтау үшін қажетті 2.24-інші корреляциялық кестені Tx динамикасы негізінде толтыру.

2.24-кесте

Жылдың реттік нөмірі (X осі)	1/x	(1/x)^2	y(Tx)	y/x
1	1,00	1,00	70,82	93,00
2	0,50	0,25	161,54	57,00
3	0,33	0,11	191,78	54,33
4	0,25	0,06	206,90	53,50
5	0,20	0,04	215,97	52,60
Сомасы	2,28	1,46	847,00	310,43

2. Тауар айналым динамикасының теңдеуінің параметрі

$$a = \frac{\sum y \sum (1/x)^2 - \sum 1/x \sum y/x}{n\sum (1/x)^2 - (\sum 1/x)^2} = 252{,}25$$

мұнда n – зерттелетін динамикалық қатардың деңгейлері.

3. Тауар айналым динамикасының теңдеуінің параметрі

$$b = \frac{n\sum y/x - \sum 1/x \sum y}{n\sum (1/x)^2 - (\sum 1/x)^2} = -181{,}43.$$

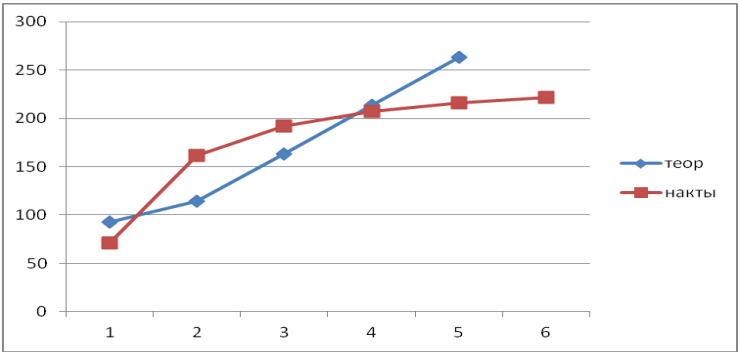
4. Талданатын кезеңдегі Тх көрсеткіштері және олардың болашағы бар кезең үшін экстраполяризацияланған мәндері, млн. тг (2.25-кесте)

2.25-кесте

2001ж: T1 = a+b/1 =	70,82115
2002ж: T1 = a+b/2 =	161,5379
2003ж: T1 = a+b/3 =	191,7768
2004ж: T1 = a+b/4 =	206,8963
2005ж: T1 = a+b/5 =	215,9679
2006ж: T1 = a+b/6 =	222,0157

5. Екі өлшемді координаттар жүйесінде нақты және теориялық тауар айналым динамикасының графиктерін тұрғызу керек, мұнда X – уақыт (жыл), Y – тауар айналым көлемі (млн. тг) (2.14-сурет).

6.



2.14-сурет. Нақты және динамикалық тауар айналымы динамикасы графигі

Нақты тауар айналым өсуге тенденциясы бар, ал теориялық тауар айналым тұрақталады.

2.6-тапсырма. 2.26-кестеде көрсетілген уақыттық қатар үшін бірінші деңгейдегі полином түрінде (сызықтық модель) трендтік модель тұрғызылған: $\hat{y}_t = 86{,}9 - 2{,}73t$. Тұрғызылған модельдің нақтылығын бағалау қажет.

2.26-кесте

t	1	2	3	4	5	6	7	8	9	10
y	90	80	77	78	69	70	63	65	56	70

2.7. тапсырма. 2010–2014 жж. ҚР-ның мата өнімін өндіру көлемі (мың кв.м) мәліметтері бойынша [37], 2.27-кестеде Қазақстан Республикасы мата өндіру көлемі (2010–2014) берілген.

2.27-кесте

Қазақстан Республикасы мата өндіру көлемі (2010–2014)

t	2010	2011	2012	2013	2014
y	36869	34122	23878	25725	41212

Келесі есептеулерді жасау қажет:

– сызықтық трендтің a_0 және a_1 коэффициенттерін және бір ай алдыға болжау жасау;

– параболалық трендтің a_0, a_1, a_3 коэффициенттеріне бір ай алдыға болжау жасау.

Бақылау сұрақтары

1. Гомперц қисығын түсінідіріңіз.
2. Өсу қисығының қандай түрлерін білесіз?
3. Өсу қисығы әдісінің артықшылықтары.

2.8. Логистикалық есептерді болжау үшін экспоненциалды тегістеу және адаптивті әдістерді пайдалану

Логистикалық есептерді болжауда адаптивті әдістерді пайдалану. Соңғы кездері бір өлшемді уақыттық қатарларды зерттеудің келешегі бар әдістерінің бірі – адаптивті әдістер және логистиканы болжау үшін пайдалану. Уақыттық қатарларды өңдеу кезінде, негізінен, соңғы кезеңнің ақпараты бағалы болады, өйткені берілген мезеттегі тенденция қалай жұмыс істейтінін білу керек. Адаптивті әдістер уақыттық қатардың деңгейлерінің әртүрлі ақпараттық құндылықтарын, мәліметтердің «ескіру» дәрежесін ескеруге мүмкіндік береді [22,30].

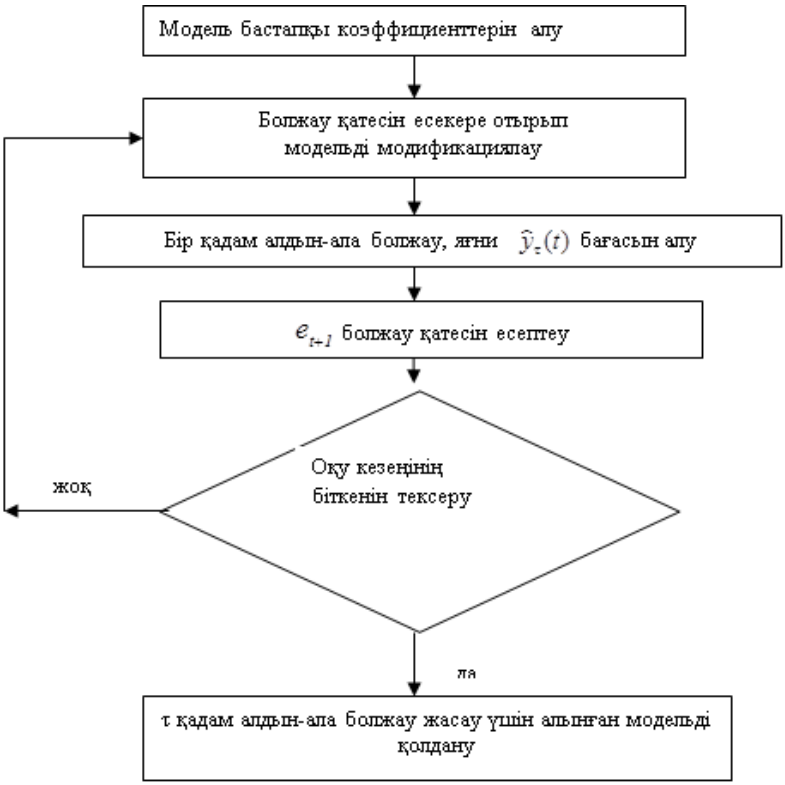
Өсу қисықтары негізінде экстраполяция әдісімен болжау жасау адаптация элементінен тұрады, өйткені "жаңа" нақты мәліметтер алынғаннан кейін қисықтардың параметрлері қайтадан есептеледі. Жаңа мәліметтердің келіп түсуі бұрын таңдалған қисықты басқа модельге ауыстыруға алып келуі мүмкін. Бірақ берілген жағдайда адаптация дәрежесі болмашы, бұған қоса ол уақыттық қатардың ұзындығы өскен сайын кемиді, өйткені мұнда әрбір жаңа нүктенің "салмағы" азаяды.

Адаптивті модельдің коэффициенттерін бағалау рекурренттік әдіс негізінде жүзеге асырылады, ол ең кіші квадраттар әдісінен, максималды ұқсастық әдісінен өзгеше болады, бұл жаңа мәліметтер пайда болған кезде есептеулердің бүкіл көлемін қайталауды қажет етпейді.

Адаптивті әдістердің артықшылықтары алдыңғы қадамда жасалынған болжау нәтижесін ескеруге қабілетті өзін-өзі түзететін модельдерді құру болып табылады. Айталық, модель коэффициенттерінің ағымды мәндері анықталған белгілі-бір күйде болсын делік. Осы модель негізінде болжау жасалады. Нақты мән келіп түскенде болжау қатесі (осы мән мен модель бойынша алынған мән арасындағы айырмашылық) бағаланылады.

Кері байланыс арқылы болжау жасау қатесі модельге келіп түседі және онда бір күйден екінші күйге өтудің қабылданған үрдісіне сәйкес ескеріледі. Нәтижесінде модельдің іс-әрекеті

қатардың динамикасына көбірек сәйкестендіру мақсатында параметрлерді түзету жасауды көздейтін "толықтырушы" өзгерістер өңделеді. Сосын уақыттың келесі мезетіне болжаудың бағасы есептелінеді, және бүкіл үрдіс қайтадан жасалады. Осылайша, адаптация қатардың әрбір нақты нүктесі алынған сайын итеративті жүзеге асырылады. Модель үнемі жаңа ақпаратты "сіңіреді", оған икемденеді және берілген мезетте жүргізіліп отырған даму тенденциясын көрсетеді. 2.15-суретте болжаудың адаптивті модельдерін құрудың жалпы сызба-нұсқасы келтірілген.



2.15-сурет. Логистикалық есептерді болжаудың адаптивті модельдерін құрудың жалпы сызба-нұсқасы

Үрдістің динамикасының өзгерісіне модельдің реакциясының жылдамдығын (жеделдігі) адаптация параметрі сипаттайды. Адаптация параметрі кездейсоқ ауытқулардың бір уақытылы фильтрациясы кезіндегі тенденциясының көрінісін қамтамасыз етілетіндей болып таңдалынуы тиіс. Адаптация параметрінің мәні эмпириялық мәндер негізінде анықталуы, аналитикалық тәсілмен шығарылуы немесе сынау әдісінің негізінде алынуы мүмкін.

Адаптация параметрін таңдау кезінде оңтайлылық критерийі ретінде болжау қатесінің орташа квадратының минимумының критерийі алынады. Қарастырылған ерекшеліктер нәтижесінде "адаптивті" деген жалпы атаумен біріктірілген болжау әдістерінің тобының анықтамасын береміз.

Адаптивті деп экономикалық-математикалық модельдерді құруға мүмкіндік беретін болжаудың өзін-өзі түзететін әдістерін атаймыз, олар алдыңғы қадамда жасалынған болжау нәтижесін ескеру және қатардың деңгейлерінің әртүрлі ақпараттық құндылықтарын есептеу жолымен шарттардың өзгерісіне оперативті жауап қайтаруға қабілетті. Берілген қасиеттерге байланысты адаптивті әдістер қысқа мерзімді болжау кезінде (бір немесе бірнеше қадам алдын ала болжау кезінде) қолданылады.

Белгілеулер:

y_t – уақыттық қатардың нақты деңгейлері;

$\widehat{y}_\tau(t)$ – t мезетінде τ уақыт бірлігіне (қадамға) алдын ала жасалынған болжам;

$e_{t+1} = y(t+1) - \hat{y}_\tau(t)$ – (t+1) нүктесінде көрсеткіштің нақты және болжаудың мәні арасындағы айырма ретінде алынған болжау қатесі.

Экспоненциалды тегістеу үрдісін қолданатын адаптивті әдістер класының ең қарапайым түрін қарастырамыз.

Логистиканы болжауда экспоненциалды тегістеу және адаптивтің модельдердін қолдану.

Қатарды экспоненциалды тегістеу үшін келесі рекуррентік формула қолданылады:

$$S_t = \alpha y_t + \beta S, \qquad (2.22)$$

мұндағы S_t – t мезетіндегі экспоненциалды орташа мәні;

α – тегістеу параметрі, α=const, 0<α<1;

β= 1-α.

Егер (2.22) қатынасын тізбектеп қолдансақ, онда S_t экспоненциалды орташаны уақыттық қатардың деңгейлерінің бұрынғы мәндері арқылы өрнектеуге болады. Мұнда $n \to \infty$ кезінде келесі өрнекті аламыз:

$$S_t = \alpha \sum_{i=0}^{n} \beta^i \cdot y_{t-1}. \qquad (2.23)$$

Осылайша, S_t шамасы қатардың барлық мүшелерінің салмақталған сомасы болып шығады. Мұнда қатардың арнайы деңгейлерінің салмақтары экспоненциалды функцияға қатысты (зерттеудің «жасына» байланысты) кемиді. Сол себепті бұл шама экспоненциалды орташа деп аталады.

Мысалы, айталық α=0,3. Онда ағымды y_t зерттеудің салмағы α=0,3, алдыңғы y_{t-1} деңгейдің салмағы α*β=0,3*0,7=0,21 болады;

y_{t-2} деңгейдің салмағы $\alpha*\beta^2$=0,147 құрайды;

y_{t-3} үшін салмақ $\alpha*\beta^3$ =0,1029 болады және т.с.с.

Айталық, уақыттық қатардың моделі $y_1=a_1+et$ болсын делік.

Ағылшын математигі Р. Браун кез келген сандық қатардың математикалық үміті және экспоненциалды орташасы сәйкес болатынын көрсетті, бірақ бұл кезде экспоненциалды орташа дисперсиясы $D[S_t]$ уақыттық қатардың дисперсиясы (σ^2)-тан кіші болатынын келтірген.

$$D[S_t] = \frac{\alpha}{2-\alpha}\sigma^2. \qquad (2.24)$$

(2.24) өрнектен жоғары α мәні кезінде экспоненциалды орташаның дисперсиясы қатардың дисперсиясынан кішігірім өзгешілігі бар екенін көрсетеді.

Осылайша, экспоненциалды орташа уақыттық қатардың тербелістерін жұтатын « фильтр» рөлін ойнай бастайды.

Бір жағынан ((2.22) сәйкес), α көбейту арқылы жаңа зерттеулер салмағын көбейту керек, екінші жағынан, кездейсоқ ауытқуларды тегістеу үшін α азайту керек.

Бұл екі талап бір-біріне қарама-қайшы болып келеді. α тегістеу параметрінің компромисті мәнін анықтау модельдің оңтайландыру есебін құрайды.

Кейде параметрдің бұл мәнін іздеу іріктеп алу арқылы жүзеге асырылады, бұл жағдайда оңтайлы ретінде қатенің ең кіші дисперсиясы алынған α мәні таңдалады. Мысалы, бұл модельдерді, " Мезозавр" пакеті көмегімен құру кезінде «оңтайландыру» бұтағы қарастырылған.

Қысқа мерзімді болжау үшін экспоненциалды орташаны қолдануда қатардың моделі келесі түрде болады деп ұйғарылған:

$$y_t = a_{1,t} + e_t,$$

мұндағы $a_{1,t}$ – қатардың деңгейінің уақыт түрін өзгертетін орташасы;

e_t – нөлдік математикалық күтімі және σ^2 дисперсиясы бар актокорреляцияланбаған ауытқулар.

Болжау модель келесі теңдікпен анықталады:

$$\hat{y}_\tau(t) = \hat{a}_{1,t},$$

мұндағы $\hat{y}_\tau$ – t мезетте жасалынған τ уақыт бірлігінің (қадам) алдын ала жасалынған болжау;

$\hat{a}_{1,t}$ – $\hat{a}_{1,t}$ бағасы, (шаманың үстіндегі ^ белгісі бағаны білдіреді).

Модельдің жалғыз $\hat{a}_{1,t}$ параметрі экспоненциалды орташамен анықталады:

$$\hat{a}_{1,t} = S_t,$$

$$\hat{a}_{1,0} = S_t.$$

(2.22) өрнегін келесі мүшелерді қайта топтап, басқаша ұсынуға болады:

$$S_t = S_{t-1} + \alpha(y_t - S_{t=1}) \qquad (2.25)$$

(y_t–S_{t-1}) шамасын болжауының қателігі ретінде қарастыруға болады. Онда жаңа болжау S_t қатесін ескере отырып, алдыңғы болжауды түзету нәтижесі ретінде алынады. Модельді адаптациялау деген осы.

Экспоненциалды тегістеу қарапайым өздігінен оқытылатын модельдің мысалы болып табылады. Мұндағы есептеулер қарапайым, итеративті орындалады, мұнда өткен мәлімет массиві жалғыз мәнге S_{t-1} азайтылған.

2.6-мысал. Айталық, ағымдағы айдағы сұраныстың болжауы 1000 дана құрайды делік. Ағымдағы айдағы нақты сұраныс 950 дана құрады. Тегістеуші константа α=0,3. Келесі айда күтілетін сұраныс келесі формуламен анықталады [33,39]:

$$y^*_{t+1} = \alpha * y_t + (1-\alpha)\cdot y^*_t;$$

y^*_{t+1} – t+1 мезетінде y^*_{t+1} болжанатын мән;

y_t – нақты мән;

y^*_t – t мезетінде y^*_t болжанатын мән.

α – салмақтық коэффициент немесе тегістеуші тұрақты [0,1].

$$y^*_{t+1} = 0.3 * 950 + (1-0.3)\cdot 1000 = 985.$$

2.7-мысал. Келесі кварталдық уақыттық қатарды бір жарым жылдағы өнімге сұраныс туралы мәліметтерді құрайды (2.28-кесте):

2.28-кесте

Өнімге сұраныс туралы мәліметтері

	Квартал			
	I	**II**	**III**	**IV**
Өткен жыл	1 200	700	900	1 100
Ағымдағы жыл	1 400	1 000	y_3 = ?	

Бізге ағымдағы жылдың үшінші кварталына болжау құру қажет. Айталық, тегістеу кезінде тұрақты α = 0,2 болсын. Алдыңғы кезеңнің болжауын алдыңғы жылдың мәліметтері

бойынша кварталдағы орташа деңгейі ретінде есептейміз. Бұдан шығатыны

y_0 =(1200+700 + 900 + 1100)/4 =975.

Өткен жылдағы сатылым болжауы орта есеппен алғанда нақты сатылымдарға сәйкес болды делік, яғни: $y_0^* = y_0 = 975$.

Онда ағымдағы жылдың бірінші кварталына жасалған болжау келесідей болады:

$$y_{0+1}^* = \alpha \cdot y_0 + (1-\alpha) \cdot y_0^* = 0.2 \cdot 975 + (1-0.2) \cdot 975 = 975.$$

Ағымдағы жылдың екінші кварталына жасалған болжау келесідей болады: $y_2^* = 0.2 \cdot 1400 + 0{,}8 \cdot 975 = 1060.$

Ағымдағы жылдың үшінші кварталына жасалған болжау келесідей болады: $y_3^* = 0.2 \cdot 1000 + 0{,}8 \cdot 1060 = 1048.$

Нәтижесінде келесі мәліметтерді аламыз (2.29-кесте):

2.29-кесте

Болжаудың нәтижелері

	Квартал			
	I	**II**	**III**	**IV**
Өткен жыл	1 200	700	900	1 100
Ағымдағы жыл	1 400	1 000	y_3 = ?	
Болжау	975	1060	1048	

Трендті ескере отырып болжауды түзету

Қарапайым экспоненциалды тегістеуді сұраныстың көбеюіне немесе азаюына тұрақты тенденция жоқ, яғни сұраныстың орташа деңгейі уақыттың жеткілікті және тұрақты болған жағдайында қолданғаны ыңғайлы. Егер, сатылымдарда, мысалы, сұраныстың көтерілуін тренд анықтаса, онда әрбір жаңа болжау нақты сұраныстан тұрақты аз болады.

Болжаудың әдістемесіне тренд бойынша есептелетін қосымша формуланы енгізе отырып түзетуге болады. Бұл үшін экспоненциалды теңдеуге трендті ескеретін тағы бір формуланы енгізу керек:

$$S_t = \alpha A_t + (1-\alpha)(S_{t-1} + T_{t-1})$$
$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1}$$
$$F_{t+1} = S_t + T_t$$

мұндағы S_t – t кезеңіндегі бастапқы болжау, T_t – t мезетіндегі тренд, F_{t+1} –трендті ескерумен t+1кезеңіндегі болжау, β – тренд үшін тегістеуші тұрақты.

2.8-мысал. Ағымдағы жылдың үшінші кварталына трендті ескере отырып, болжауды есептейік (2.30-кесте):

2.30-кесте

Кварталдар бойынша сұраныс туралы мәліметтер

	Квартал			
	I	**II**	**III**	**IV**
Өткен жыл	1 200	700	900	1 100
Ағымдағы жыл	1 400	1 000	F_3 = ?	

Алдымен ағымдағы жылдың бірінші кварталының болжауын есептейік. Есептеу кезінде кіріс мәліметтері ретінде S_0 = 975 (алдыңғы жылдың мәліметтері бойынша кварталдағы орташа сұраныс) және T_0 = 0 (тренд жоқ) қолданамыз. Тегістеуші тұрақтылар үшін α = 0,2 және β = 0,3 болсын делік. Енді есептеулерді бастаймыз.

Ағымдағы жылдың бірінші кварталына жасалған болжау келесідей болады:

$S_0 = 975$, $T_0 = 0 \rightarrow F_1 = 975 + 0 = 975$.

Ағымдағы жылдың екінші кварталына жасалған болжау келесідей болады:

$S_1 = 0{,}2\times1400 + 0{,}8\times(975 + 0) = 1060$;

$T_1 = 0{,}3\times(1\,060 - 975) + 0{,}7\times0 = 25{,}5$;

$F_2 = 1060 + 25{,}5 = 1085{,}5$.

Ағымдағы жылдың үшінші кварталына жасалған болжау келесідей болады:

$S_2 = 0{,}2\times1000 + 0{,}8\times(1060 + 25{,}5) = 1068{,}4$;

$T_2 = 0{,}3\times(1068{,}4 - 1060) + 0{,}7\times25{,}5 = 20{,}37$;

$F_2 = 1068{,}4 + 20{,}37 = 1088{,}77$.

Нәтижесінде алатынымыз (2.31-кесте):

2.31-кесте

Болжаудың нәтижелері

	Квартал			
	I	**II**	**III**	**IV**
Өткен жыл	1 200	700	900	1 100
Ағымдағы жыл	1 400	1 000		
Болжау	975	1 086	1 089	

Болжауды трендті және маусымдылықты ескере отырып түзету

Болжау кезінде трендті ғана емес, сұраныстың маусымдық тербелістерін де ескеруге болады. Келесі мысалда қарастырылған модельді қарастырудан бұрын, уақыттық қатарды келесі шарттардың орындалуына тексеру керек:

1. Сұраныстың маусымдық шыңдары мен түсулері статистикалық қатарда айқын көрінуі керек, яғни олар сұраныстың кездейсоқ тербелістерінен көбірек болу керек.

2. Маусымдық шығындары мен түсімдері тұрақты түрде жылдан жылға қайталанылуы тиіс.

Егер бұл екі шарт орындалмаса, онда маусымдық тербелістер тұрақсыз, білінер-білінбес және «шудан» ажыратылуы қиын болады, бұл жағдайда келесі уақыттық кезеңге сұранысты дәл болжау үшін қолдану қиын болады. Егер шарттар орындалса және сұраныс тербелістерінің үлкен амплитудасын ескеру үшін модельде тегістеуші тұрақтының жоғарғы мәні орнатылса, онда модельді күрделендірудің мәні бар [22,27,33].

Бұл жаңа модельде индекстер түрінде берілетін трендтің және маусымдылықтың түзетулерін ескере отырып құрылады. Бұл болжаудың жоғары дәлдігіне жетуге мүмкіндік береді.

Күрделендірілген модельдің теңдеулері:

$$S_t = \alpha(A_t / I_{t-L}) + (1-\alpha)(S_{t-1} + T_{t-1})$$

$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1}$$

$$I_t = \gamma\left(A_t / S_t\right) + (1-\gamma) I_{t-L}$$

$$F_{t+1} = (S_t + T_t) I_{t-L+1}$$

мұндағы T_t – t кезеңіндегі тренд, S_t – t кезеңіндегі бастапқы болжау, F_{t+1} – трендті және маусымдылықты ескере отырып жасалынған t+1 кезеңіндегі болжау, I_t – t кезеңіндегі маусымдық тербелістер индексі, L – толық маусымдық цикл өтетін уақыттық кезең, γ – маусымдық индекс үшін тегістеуші тұрақты.

2.9-мысал. Трендті ескере отырып ағымдағы жылдың үшінші кварталына жасалынған болжауды есептейміз (2.32-кесте):

2.32-кесте

Өнімге сұраныс туралы мәліметтер

	Квартал			
	I	**II**	**III**	**IV**
Өткен жыл	1 200	700	900	1 100
Ағымдағы жыл	1 400	1 000	F_3 = ?	

Алдымен ағымдағы жылдың бірінші кварталына жасалынған болжауды есептейміз. Есептеу кезінде кіріс мәндері ретінде S_{t-1}=975 (алдыңғы жылдардың мәліметтері бойынша кварталдағы орташа сұраныс) және T_{t-1} = 0 (тренд жоқ) аламыз. Мұнда тегістеуші тұрақты ретінде $\alpha = 0{,}2$ және $\beta = 0{,}3$, ал $\gamma = 0{,}4$ болсын делік. Енді есептеулерді бастаймыз.

Ағымдағы жылдың бірінші кварталына жасалынған болжау келесідей болады:

$S_0 = 975$ и $T_0 = 0$. Онда:

$F_1 = (975 + 0) \times 1{,}23 = 1200$, өйткені $I_1 = 1200 / 975 = 1{,}23$.

Онда:

Ағымдағы жылдың екінші кварталына жасалынған болжау келесідей болады:

$S_1 = 0{,}2 \times 1400 / 1{,}23 + 0{,}8 \times (975 + 0) = 1007{,}5$;

$I_1 = 0{,}4 \times 1400 / 1007{,}5 + 0{,}6 \times 1{,}23 = 1{,}29$;

$T_1 = 0{,}3 \times (1007{,}5 - 975) + 0{,}7 \times 0 = 9{,}75$;

$F_2 = (1007{,}5 + 9{,}75) \times 0{,}72 = 730{,}3$, өйткені $I_2 = 700 / 975 = 0{,}72$.

Ағымдағы жылдың үшінші кварталына жасалынған болжау келесідей болады:

$S_2 = 0,2 \times 1000 / 0,72 + 0,8 \times (1007,5 + 9,75) = 1092,4;$

$I_2 = 0,4 \times 1000 / 1092,4 + 0,6 \times 0,72 = 0,8;$

$T_2 = 0,3 \times (1092,4 - 1007,5) + 0,7 \times 9,75 = 32,3;$

$F_2 = (1092,4 + 32,3) \times 0,92 = 1005$, өйткені $I_3 = 900 / 975 = 0,92$.

Нәтижесінде алатынымыз:

2.33-кесте

Болжаудың нәтижелері

	Квартал			
	I	**II**	**III**	**IV**
Өткен жыл	1 200	700	900	1 100
Ағымдағы жыл	1 400	1 000		
Болжау	1 200	779	1 005	

*2.*10-м*ысал*. Біздің мысалымызға қайта ораламыз. Экспоненциалды тегістеудің базалық моделінде келесі нәтижелер алынған еді (2.34-кесте):

2.34-кесте

Нақты және болжаудың мәндер туралы мәліметтер

	Квартал			
	I	**II**	**III**	**IV**
Өткен жыл	1 200	700	900	1 100
Ағымдағы жыл	1 400	1 000	$F_3 = ?$	
Болжау	1 200	779	1 005	

Бізге нақты және болжау мәндері белгілі ағымдағы жылдың бірінші және екінші кварталдың мәліметтері бойынша болжаудың стандартты қатесін бағалаймыз. Айталық, сұранысты болжауға қатысты нормальді үлестірілген делік. Үшінші квартал үшін 95 % ықтималдығы бар сенімділік интервалының шекараларын есептейміз.

Болжаудың стандартты қатесі:

$$S_E = \sqrt{\frac{(1400-1200)^2 + (1000-779)^2}{2-1}} = 298.$$

А кестесін қолдана отырып, (қосымша I қара), $z_{95\%}$ = 1,96 коэффициентін анықтаймыз және сенімділік интервалының шекараларын келесі формула бойынша аламыз:

$$Y = F_3 \pm z(S_E) = 1005 \pm 1{,}96 \times 298 = 1064 \pm 584{,}2.$$

Бұдан шығатыны, ағымдағы жылдың үшінші кварталына сұраныстың сенімділік интервалы шекараларының 95 % ықтималдығы бар болжауы келесідей болады: 420,8 < Y < 1589,2.

2.8-тапсырма. Браун сызықтық моделі бойынша болжау жасау.

Жаңа тауардың сатылым көлемі туралы 35 апта ішіндегі қолда бар ақпарат бойынша (2.35-кестедегі y) сызықтық тенденциясы бар Брауннның адаптивті моделін тұрғызу. 0,3 тегістеу параметрін қолданып, үш қадам алдыға болжау жасау. Модельдеу және болжау нәтижелерін графикте көрсету.

2.35-кесте

Бастапқы мәліметтер

T	1	2	3	4	5	6	7
Y	28	43	45	65	71	54	72
T	8	9	10	11	12	13	14
Y	75	102	112	126	190	165	193
T	15	16	17	18	19	20	21
Y	204	221	268	264	280	322	318
T	22	23	24	25	26	27	28
Y	342	351	368	390	380	400	410
T	29	30	31	32	33	34	35
Y	427	402	400	425	435	395	440

Бақылау сұрақтары

1. Экспоненциалды тегістеу әдісінің артықшылықтары мен кемшіліктерін көрсетіңіз.
2. Экстраполяция әдісінің артықшылығы неде?
3. Болжау жасаудың адаптивті модельдерді құру алгоритмін түсіндіріңіз.

2.9. Жоғары реттегі адаптивті полиномиалды модельдер

Экспоненциалды орташа түсінігін жоғары реттегі экспоненциалды орташа жағдайында жалпылауға болады.

р-інші реттегі тегістеу:

$$S_t^{(p)} = \alpha \cdot S_t^{(p-1)} + \beta \cdot S_{t-1}^{(p)}. \qquad (2.26)$$

Бұл (*p-1)*-інші реттегі тегістеу нәтижелеріне қолданылатын қарапайым экспоненциалды тегістеу болып табылады [22,30].

Егер әйтеуір бір үрдістің тренді n дәрежелі полиноммен сипатталатын болса, онда болжанатын полиномның коэффициенттері тиісті реттегі экспоненциалды орташалар арқылы есептеді.

Детерминацияланған және кездейсоқ құрамадан тұратын зерттелетін үрдіс n-інші реттегі полиноммен сипатталса, r қадам алдын ала жасалатын болжау келесі формуламен жүзеге асырылады:

$$\hat{y}_\tau(t) = \hat{a}_1 + \hat{a}_2\tau + \frac{1}{2}\hat{a}_3\tau^2 + ... + \frac{1}{n\text{I}}\hat{a}_{n+1}\cdot\tau^n, \qquad (2.27)$$

мұндағы $\hat{a}_1, \hat{a}_2, \hat{a}_3, ..., \hat{a}_{n+1}$ – параметрлердің бағалаулары.

Алғаш рет Р. Браун және Р. Майермен дәлелденген экспоненциалды тегістеу және болжау әдісінің фундаменталді теоремасы n-інші реттегі полиномның белгісіз (n+1) коэффициенттері $\hat{a}_1, \hat{a}_2, \hat{a}_3, ..., \hat{a}_{n+1}$, экспоненциалды орташа $S_t^{(i)}$ сызықты комбинациялар көмегімен бағаланылады, мұнда $i = \frac{1}{n+1}$.

Бұдан шығатыны, есеп реті 1-ден n+1-ге дейін өзгеретін экспоненциалды орташаларды есептеуге, сосын олардың сызықтық комбинациялары арқылы полином коэффициенттерін есептеуге ұштасады.

Іс жүзінде екінші реттен жоғары полиномдар қолданылады. Мысалы, бірінші реттегі полиномды қолдану кезінде уақыттық

қатардың бірінші реттегі адаптивті моделінің түрі келесідей болады:

$$y_t = a_{1,t} + a_{2,t} + e_t\,, \qquad (2.28)$$

мұндағы $a_{1,t}$ – ағымды t-інші деңгейдің мәні;
$a_{2,t}$ – ағымдағы өсімнің мәні.

(2.36)-кестеде бұл модельдер бойынша есептеулер үшін қажетті формулалар келтірілген.

2.36-кесте

Адаптивті полиномиалды модельдер бойынша болжау жасау үшін арналған негізгі формулалар

Модель дәрежесі	Бастапқы шарт	Экспоненциалды орташа	Коэффициенттердің бағасы	Болжау моделі
n=0	$s_0^{(1)} = \hat{a}_{1,0}$	$s_t^{(t)} = ay_t + \beta s_{t=1}^{(1)}$	$\hat{a}_{1,t} = s_t^{(1)}$	$\hat{y}_\tau(t) = \hat{a}_{1,t}$
n=1	$s_0^{(1)} = \hat{a}_{1,0} - \frac{\beta}{\alpha}\hat{a}_{2,0}$ $s_0^{(1)} = \hat{a}_{1,0} - \frac{2\beta}{\alpha}\hat{a}_{2,0}$	$s_t^{(t)} = ay_1 + \beta s_{t-1}^{(1)}$ $s_t^{(2)} = aS_t^{(1)} + \beta s_{t-1}^{(2)}$	$\hat{a}_{1,t} = 2S_t^{(1)} - S_t^{(2)}$ $\hat{a}_{1,t} = \frac{\alpha}{\beta}(S_t^{(1)} - S_t^{(2)})$	$\hat{y}_\tau(t) = \hat{a}_{1,t} +$ $+ \hat{a}_{2,t}\cdot\tau$
n=2	$s_0^{(1)} = \hat{a}_{1,0} - \frac{\beta}{\alpha}\hat{a}_{2,0} +$ $+ \frac{\beta(2-\alpha)}{2\alpha^2}\hat{a}_{3,0}$ $s_0^{(2)} = \hat{a}_{1,0} - \frac{2\beta}{\alpha}\hat{a}_{2,0} +$ $+ \frac{\beta(3-2\alpha)}{\alpha^2}\hat{a}_{3,0}$ $s_0^{(3)} = \hat{a}_{1,0} - \frac{3\beta}{\alpha}\hat{a}_{2,0} +$ $+ \frac{3\beta(4-3\alpha)}{2\alpha^2}\hat{a}_{3,0}$	$s_t^{(1)} = \alpha y_1 + \beta S_{t-1}^{(1)}$ $s_t^{(2)} = \alpha S_t^{(1)} + \beta S_{t-1}^{(2)}$ $s_t^{(3)} = \alpha S_t^{(2)} +$ $+ \beta S_{t-1}^{(3)}$	$\hat{a}_{1,t} = 3(S_t^{(1)} - S_t^{(2)} + S_t^3)$ $\hat{a}_{2,t} = \frac{\alpha}{2\beta^2}\left[(6-5\alpha)S_t^{(1)}\right] -$ $-2(5-4\alpha)S_t^{(2)} + (4-3\alpha)S_t^{(3)}$ $\hat{a}_{3,t} = \frac{\alpha}{\beta^2}$ $\left[S_t^{(1)} - 2S_t^{(2)} + S_t^{(3)}\right]$	$\hat{y}_\tau(t) = \hat{a}_{1,t} +$ $+ \tau\cdot\hat{a}_{2,t} + \frac{1}{2}\tau^2\hat{a}_{3,t}$

Экспоненциалды тегістеу әдісі бойынша уақыттық қатарларды болжау үрдісі өте қарапайым және ол келесі кезеңдерден тұрады:

1. Экспоненциалды тегістеу моделінің түрі таңдалады, *a* тегістеу параметрінің мәні беріледі. Адаптивті полиноминалды модельдің ретін таңдау кезінде әртүрлі әдістемелер, мысалы, графикалық талдау, айырмашылық өзгерісінің әдісі қолданылады.

2. Бастапқы шарттар анықталады. Мысалы, бірінші реттегі полиномиалды моделі үшін 1,0; 2,0 анықтау керек.

Көбінесе бұл бағалар ретінде ең кіші квадраттар әдісімен алынған тиісті полиномдардың коэффициенттері алынады. Нөлдік реттегі модель үшін бастапқы шарттар қатардың бірінші бірнеше теңдеулерін орташаландыру арқылы есептеледі. Бұл бағаларды біле отырып, кестеде көрсетілген формулалар арқылы экспоненциалды орташалардың бастапқы мәндері анықталады.

3. Тиісті экспоненциалды орташалардың мәндері анықталады.

4. Модельдің коэффициенттерінің бағалары табылады.

5. Бір нүктеге алдын ала болжау жүзеге асырылады, уақыттық қатардың нақты мәнінің болжанылған мәнінен ауытқуы анықталады. Берілген үрдістің 3-тен 5-ке дейінгі қадамдары барлық t<=n болғанша қайталанылады, мұнда n – қатардың ұзындығы.

6. Ақырғы болжау моделі t=n мезетінде соңғы қадамда қалыптасады. Болжау (2.27) өрнегі негізінде оған t артықшылықпен сатып алу уақытын және соңғы коэффициенттер мәндерін қою арқылы алынады.

Қарастырылған модельдердің оң ерекшеліктеріне жаңа ақпарат келіп түскенде есептеулерді қайтадан қайталау қажет болмайтындығын жатқызуға болады. Бастапқы шарт ретінде $S_t^{(i)}$ тегістеу функциясының соңғы мәндерін алған жеткілікті болады және есептеулерді жалғастыру керек.

Маусымдық тербелістер үшін адаптивті модельдер.

Көптеген экономикалық уақыттық қатарлар кезеңдік маусымдық тербелістерден тұрады. Мұндай қатарлар маусымдылықтың мультипликативті (2.29.) және аддитивті коэффициенттерімен (2.30) екі типті модельдермен сипатталады:

$$y_t = a_{1,t} \cdot f_t + e_t$$

$$y_t = a_{1,t} + g_t + e_t$$

мұндағы $a_{1,t}$ – дамудың тенденциясының сипаттамасы;

$g_t, g_{t-1}, ..., g_{t-e+1}$ – маусымдылықтың аддитивті коэффициенттері;

$f_t, f_{t-1}, ..., g_{r-e+1}$ – маусымдылықтың мультипликативті коэффициенттері;

e – толық маусымдық циклдегі фазалардың мөлшері (ай сайынғы зерттеулер үшін e=12, кварталдық -e = 4), e_t – нөлдік математикалық күтімі бар кездейсоқ құрама.

Аддитивті және мультипликативті маусымдық әсерін үйлестіре отырып, тенденциялардың әртүрлі комбинацияларын тере отырып, адаптивті маусымдық модельдерді құруға болатыны анық. Әйтеуір бір модельді таңдау зерттелетін үрдістің динамикасының сипатына байланысты болады. Мысал ретінде тенденциясының сипаты сызықты және мультипликативті маусымдық әсері бар Уинтерс моделін қарастырамыз. Бұл модель Хольтің сызықтық өсімінің екі параметрлі моделі мен Уинтерс маусымдық моделінің бірлігі болып табылады. Сондықтан оны көбінесе Хольт-Уинтерс моделі деп атайды.

Хольт-Уинтерс моделі бойынша τ қадам алдын ала болжау келесі өрнекпен анықталады:

$$\hat{y}_t(t) = \left(\hat{a}_{1,t} + \tau \cdot \hat{a}_{2,t}\right) \cdot \hat{f}_{t-e+\tau}. \qquad (2.31)$$

Коэффициенттердің жаңартылуы келесі түрде жаңартылады:

$$\hat{a}_{1,t} = \alpha_1 \frac{y_t}{\hat{f}_{t-e}} + (1-\alpha_1)\left(\hat{a}_{1,t-1} + \hat{a}_{2,t-1}\right)$$

$$\hat{f}_t = \alpha_2 \frac{y_t}{\hat{a}_{1,t}} + (1-\alpha_2)\hat{f}_{t-e}$$

$$\hat{a}_{2,t} = \alpha_3\left(\hat{a}_{1,t} - \hat{a}_{1,t-1}\right) + (1-\alpha_3)\hat{a}_{2,t-1}$$

$$0 \prec \alpha_1, \alpha_2, \alpha_3 \prec 1$$

(2.30) өрнектен $\widehat{a}_{1,t}$ маусымдық тербелістерден y_t нақты мәліметтерді және алдыңғы $\widehat{a}_{1,t-1}$ бағаны тазарту жолымен алынған ағымдағы $\frac{y_t}{\widehat{f}_{t-e}}$ бағасының салмақталған сомасы болып табылады. Маусымдық f_t коэффициенті ретінде ($\widehat{f}_{t-e}$) циклінің осындай фазасы үшін жасалынған ең кеш баға алынады.

Сосын бірінші теңдеу бойынша алынған $\widehat{a}_{1,t}$ шамасы екінші теңдеу жаңа маусымдық коэффициентін анықтау үшін қолданылады. $\widehat{a}_{2,t}$ бағалары экспоненциалды тегістеуге ұқсас үрдіс бойынша модифицияланылады.

α_1, α_2, α_3 үшін оңтайлы мәндерді Уинтерс бұл параметрлердің мәндерінің торын бере бере отырып, тәжірибелік жолмен табуды ұсынады. Мұнда қатенің стандартты ауытқуы салыстыру критерийі болады.

Адаптивті маусымдық модельдер болжау есептерін шешуге бағытталған қолданбалы бағдарламалардың қазіргі пакетінің маңызды құрамасы болып табылады.

2.11-мысал. Қоймадағы ағымды қорды болжау.

Жұмыста алынған қоймадағы детальдардың жұмсалу мәліметтері негізінде болжау әдістерінің қолданылуын қарастырайық [31]. 2.37-кестеде ағымды қордың үш түрлі жұмсалуы келтірілген, әрбір жұмсалу үшін тиісті циклдегі қоймадан детальдардың жұмсалуын көрсететін бір күндегі шығын шамалары және интегралды сипаттамалар келтірілген.

Бір жұмсалу және үш жұмсалу ансамблі үшін болжаудың мүмкін нұсқаларын көрсетейік.

Бірінші жұмсалуды қарастырайық. Айталық, бізге жұмыстың бес күнінің детальдарды қолдану мәндері белгілі делік, 2.38-кесте.

Сызықтық тәуелділік түріндегі $\overline{y_t}$ тренд теңдеуін таңдаймыз

$$y_t = a_0 + a_1 t.$$

Теңдеудің a_0 және a_1 коэффициенттері келесі формулалармен анықталады:

$$a_0 = \frac{\sum y_i \sum t_i^2 - \sum t_i \sum y_i t_i}{N\sum t_i^2 - (\sum t_i)^2};$$

$$a_1 = \frac{N\sum y_i t_i - \sum y_i \sum t_i}{N\sum t_i^2 - (\sum t_i)^2}.$$

(2.28.),(2.29) формулалары ең кіші квадраттар әдісі көмегімен алынғанын еске түсірейік.

Формулаға кіретін сомалар мәндері 2.37-кестеде есептелген. Олардың мәндерін қолдана отырып, a_0 = 45,2, a_1 = -3,0 анықтаймыз. Осылайша, болжау теңдеуі келесі түрде болады:

$$y_t = 45{,}2 - 3{,}0t.$$

2.37-кесте

Қорды жұмсаудың үш циклі ішіндегі сұраныс динамикасы

1-інші цикл			2-інші цикл			3-інші цикл		
Күн j	Сұраныс, бірл.	Цикл басынан барлығы	Күн j	Сұраныс, бірл.	Цикл басынан барлығы	Күн j	Сұраныс, бірл.	Цикл басынан барлығы
1	9	9	11	0	0	21	5	5
2	2	11	12	6	6	22	5	10
3	1	12	13	5	11	23	4	14
4	3	15	14	7	18	24	3	17
5	7	22	15	10	28	25	4	21
6	5	27	16	7	35	26	1	22
7	4	31	17	6	41	27	2	24
8	8	39	18	9	50	28	8	32
9	6	45	19	*	50	29	3	35
10	5	50	20	*	50	30	4	39
Ескерту : * – тапшылық								

Интервалды болжаудың шегін анықтау үшін орташа квадраттық ауытқуды есептеу керек:

$$\sigma_i = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(y_i - \overline{y_t})^2}.$$

Қосымша есептеулер 2.38-кестеде келтірілген. (2.30) формулаға мәндерді қойып, σ_t анықтаймыз.

$$\sigma_t = \sqrt{\frac{1}{5-1} \cdot 13} = 1{,}8.$$

2.38-кесте

Кіріс мәліметтер және N = 5 кезінде теңдеу коэффициенттерін есептеу нәтижелері

t_i, *күн*	y_i, *бірл.*	t_i^2	$y_i t_i$	*болжау* y_t^*	$(y_t - y_i)^2$
1	41	1	41	42	1
2	39	4	78	39	0
3	38	9	114	36	4
4	35	16	140	33	4 =1
5	28	25	140	30	4
$\sum t_i = 15$	$\sum y_i = 181$	$\sum t_i^2 = 55$	$\sum_{y_i}^{t} = 513$		$\sum (y_t - y_i)^2 = 13$
∗Ескерту: мәндер жуықталған					

2.38-кестеде ағымды шығынның үш жұмсалуы келтірілген; әрбір жұмсалу үшін бір күндегі шығын мөлшері және интегралды сипаттамалар келтірілген.

Бірінші жұмсалудың 5 күндегі сұраныс мәндерін қолданамыз. Алдымен α = 0,4 деп аламыз. Бастапқы шарт ретінде бірінші күндегі сұраныс мәнін y^*_t =9 таңдаймыз. t=1 кезінде болжам жасаймыз, формулаға қажетті мәндерді және бастапқы шартты қоямыз:

$$y^*_{1+1} = 0{,}4 \bullet 9 + (1 - 0{,}4) \bullet 9 = 9 \textit{бірлік}.$$

Екінші кезеңдегі болжам бастапқы шартқа тең. Екінші деңгей үшін болжам бастапқы шартқа тең. Екінші деңгей үшін нақты мән y_2 =2. t=2 үшін экспоненциалды орташаның мәнін анықтаймыз:

$$y^*_{2+1} = 0{,}4 \bullet 2 + (1 - 0{,}4) \bullet 9 = 6{,}2{,}3 \textit{бірл}.$$

Үшінші күндегі болжам 6,2 бірлік.

t=3 кезіндегі экспоненциалды орташа және төртінші күнге болжам келесідей болады:

$$y^*_{3+1} = 0{,}4 \bullet 1 + (1 - 0{,}4) \bullet 6{,}2 = 4{,}12 \textit{бірлік}.$$

Осылайша, бесінші және алтыншы күнге болжам жасалады:

$$y^*_{4+1} = 0{,}4 \bullet 3 + (1 - 0{,}4) \bullet 4{,}12 = 3{,}67 \textit{бірлік},$$

$$y^*_{5+1} = 0{,}4 \bullet 7 + (1 \quad 0{,}4) \bullet 3{,}67 = 5{,}00 \text{ бірлік}.$$

Сонымен, алтыншы күнге жасалған болжам 5 бірлік құрайды. Енді болжам қатесін анықтаймыз:

$$s = \sqrt{\frac{(2-9)^2 + (1-6{,}2)^2 + (3-4{,}12)^2 + (7-3{,}67)^2}{5-1}} = 4{,}7 \text{ бірлік}.$$

Интервалды болжау жасаймыз. Кестеде 0,1; 0,05 және 0,01 мәнділік деңгейі үшін Стьюдент критерийінің мәні келтірілген.

Айталық, мәнділік деңгейі 0,1-ге тең. Болжаудың төменгі шегі

$$y_m = 5 - 4{,}7 \bullet 2{,}132 = -5{,}02.$$

Төменгі шекті нольге тең деп аламыз. Болжамның жоғарғы шегі

$$y_{жо} = 5 + 4{,}7 \circ 2{,}132 = 11{,}83.$$

Осылайша, экспоненциалды реттеу әдісі көмегімен алтыншы күнге болжам жасалды: сұраныстың орташа 5 бірлік, болжам қатесі 4,7 бірлік, 0,9 ықтималдықпен 0-ден 12 бірлік интервалында сұраныс күтілуде.

2.9-тапсырма. 2.11-мысалдан y_i және σ_t алынған тәуелділіктер негізінде болжау бағаларын анықтаңдар:

1. $\bar{T}$ ағымды қорды жұмсаудың орташа уақыты;
2. Р сенімділік интервалымен y_c сақтандыру қоры;
3. Болжау жасалатын кезең ішінде қоймадағы тапшылық жоқ болу ықтималдығы.

2.10-тапсырма. 2.11-мысалдағы мәліметтердің үшінші жүзеге асырылуды қарастырыңыз және бірінші күндегі сұранысқа, яғни 5 бірлікке тең бастапқы шарттармен және α = 0,4 параметрі кезінде экспоненциалды тегістеу әдісімен алтыншы күнге болжау жасаңыз.

2.11-тапсырма. Мемлекеттердің кез келгені үшін 2006–2010 жж. бойынша сатылым көлемі бойынша ай сайынғы мәліметтерді қолданып регрессиялық, авторегрессиялық және Хольта-Винтер жылға болжау жасаңыз. Тегістеу коэффициенті 0,5-ке тең.

2.39-кесте

Бастапқы мәліметтер

Жылдар 2006	Ай	Сатылым, Тг	2007	Ай	Сатылым, Тг	2008	Ай	Сатылым, Тг	2009	Ай	Сатылым, Тг
1	1,00	17 986 229	1	1,00	21 149 853	1	1,00	25 379 305	1	1,00	24 255 101
2	2,00	23 571 965	2	2,00	23 770 186	2	2,00	27 823 570	2	2,00	22 391 876
3	3,00	25 537 589	3	3,00	29 608 386	3	3,00	28 518 039	3	3,00	27 902 911
4	4,00	24 630 951	4	4,00	28 588 548	4	4,00	33 971 886	4	4,00	24 102 028
5	5,00	24 429 696	5	5,00	29 712 036	5	5,00	31 577 081	5	5,00	24 939 643
6	6,00	26 116 377	6	6,00	31 191 793	6	6,00	29 328 611	6	6,00	25 401 741
7	7,00	27 931 501	7	7,00	28 311 730	7	7,00	34 312 920	7	7,00	22 817 314
8	8,00	25 914 893	8	8,00	27 438 262	8	8,00	31 364 478	8	8,00	23 554 471
9	9,00	24 904 130	9	9,00	26 166 319	9	9,00	29 046 432	9	9,00	21 219 769
10	10,00	22 360 354	10	10,00	25 916 207	10	10,00	27 244 171	10	10,00	21 144 903
11	11,00	23 825 299	11	11,00	23 168 086	11	11,00	24 353 446	11	11,00	19 185 427
12	12,00	22 241 744	12	12,00	27 707 909	12	12,00	25 447 525	12	12,00	20 507 490

Бақылау сұрақтары

1. Динамика қатарының қандай модельдерін білесіз?
2. Хольта-Уинтерс моделі бойынша болжау қалай жасалады?
3. Мауысымдық ауытқулардың адаптивті моделінің мультипликативті моделінен ерекшелігі неде?
4. Уақыттық қатар үшін экспоненциалды орташа қалай есептеледі?

БИБЛИОГРАФИЯЛЫҚ ТІЗІМ

1. *Мухамадиев Б. М.* Эконометрика и эконометрическое прогнозирование: Учебное пособие. – Алматы: «Қазақ университеті», 2007. – 250 с.

2. *Абезгауз Г. Г., Тронь А. П., Копенкин Ю. Н., Коровнина И. А.* Справочник по вероятностным расчетам. Военное издательство МО СССР. – М.,1970. – 527 с.

3. *Доугерти К.* Введение в эконометрику. – М.: ИНФРА -М, 1999. – 402 с.

4. *Новиков А. И.* Эконометрика: Учебное пособие. – М.: ИНФРА-М, 2006.

5. Практикум по эконометрике: Учебник /Под ред. И. И. Елесеевой. – М.: Финансы и статистика, 2002. – 215 с.

6. *Елесеева И. И., Курышева С. В. и др.* Эконометрика: Учебник. – М.: Финансы и статистика, 2005. – 576 с.

7. *Исмаилова Р. Т.* и др. Элементы математической статистики. Методические указания по практическим занятиям для всех экономических специальностей по Эконометрике. – Алматы: КазНТУ, 2006.

8. *Алимжанова Р. М. Исмаилова Р. Т.* Основы эконометрики, Методические указания по практическим занятиям для всех экономических специальностей по Эконометрике. – Алматы: КазНТУ, 2006.

9. *Исмаилова Р.Т.* Эконометрика: Учебное пособие. – Алматы: КазНТУ имени К. И. Сатпаева, 2010. – 156 с.

10. *Алимжанова Р. Т., Казбекова К. К., Исмаилова Р. Т.* «Эконометрика» пәні бойынша практикалық сабақтарды жүргізу және студенттердің өзіндік жұмыстарын орындауға арналған әдістемелік нұсқау («Экономика» 050506, 050507, 050508, 050509, 050510, 050511 мамандқтарының студенттері үшін). − Алматы: ҚазҰТУ, 2010. – 25 б.

11. *Дильдебаева Ж. Т.* Эконометрика: Оқу құралы. – Алматы: ҚазҰТУ, 2011. – 152 б.

12. Эконометрика: Учебно-методическое пособие / *Шалабанов А. К., Роганов Д. А.* – Казань: ТИСБИ, 2002. – 56 с.

13. *Вентцель Е. С.* Теория вероятностей. 3-е изд. – М.: Наука. 1998.

14. *Гмурман В. Е.* Руководства к решению задач по теории вероятностей и математической статистике. Уч.пос. для вузов. – М.: Высшая школа, 1999.

15. *Дорохина Е. Ю., Переснякова Л. Ф., Тихомиров Н. П.* Сборник задач по эконометрике: Учебное пособие. – М.: Изд-во «Экзамен», 2003.

16. *Магнус Я. Р, Катышев П. К, Пересецкий А. А.* Эконометрика. Начальный курс. – М.: Дело, 1997. – 248 с.

17. *Казбекова К. К.* Статистика: Оқу құралы. – Алматы: ҚазҰТУ, 2009. –142 б.

18. *Кендэл М.* Временные ряды. – М.: "Финансы и статистика", 1981.

19. *Кильдишев Г. С., Френкель А. А.* Анализ временных рядов и прогнозирование. – М.: "Статистика", 1973.

20. *Скучалина Л. Н., Крутова Т. А.* Организация и ведение базы данных временных рядов. Система показателей, методы определения, оценки прогнозирования информационных процессов, ГКС РФ. – М., 1995.

21. Статистическое моделирование и прогнозирование: Учебное пособие. (Под ред. А. Г. Гранберга). – М.: "Финансы и статистика", 1990.

22. *Маленво Э.* Статистические методы эконометрики. Перевод с французского А. И. Гладышевского, Г. А. Фреймана. – М.: Статистика, 1976.

23. *Четыркин Е. Н.* Статистические методы прогнозирования. – М.: "Статистика", 1975.

24. *Френкель А. А.* Прогнозирование производительности труда: методы и модели. – М.: "Экономика", 1989.

25. Экономико-математические методы и прикладные модели. (Под ред. В. В. Федосеева). – М.: «Юнити», 1999.

26. *Александренков Ю. В.* Логистика: практикум. – Минск. БНТУ, 2010. – 61 с.

27. *Гаджинский А. М.* Логистика: Учебник для студентов высших учебных заведений. – 12-е изд., перераб. и доп. – М: Издательско-торговая корпорация «Дашков и К°», 2006. – 432 с.

28. *Гаджинский А. М.* Практикум по логистике (учебник). – М.: ИКЦ Маркетинг, 2001. 11 п.л.

29. *Герасимов Борис Иванович.* Основы логистики: Учеб. пособие для сред. проф. образования / Герасимов Борис Иванович, Жариков В. В., Жариков В. Д. – М.: ФОРУМ, 2010. – 304 с.

30. *Лукашин Ю. П.* Адаптивные методы краткосрочного прогнозирования. – М.,“Статистика”, 2010.

31. *Лукинский В. С.* Модели, методы теории логистики: Уч.пособие, 2-е издание. Издательство: Питер, 2008. – 448 с.

32. *Б. К. Плоткин., Л. А. Делюкин.* Экономико-математические методы и модели в коммерческой деятельности и логистике. Сборник задач и упражнений: Учебное пособие. – СПб: Изд-ство СПбГЭУ, 2013. – 64 с.

33. http://www.aicherkesov. Логистика для МВА. Прогнозирование спроса.

34. *Кулжабай Н.М., Исмаилова Р.Т.* Теория организационного управления логистическими системами: Учебное пособие. – Алматы: КазНТУ им. К. И. Сатпаева, 2013. – 146 с.

35. *Кулжабай Н. М., Тышканбаева М. Б., Исмаилова Р. Т., Ботаева С.* Логистика: Учебное пособие. – Алматы: КазНТУ, 2014. – 289 с.

36. *Исмаилова Р. Т., Оразымбетова А. Ш., Абдикул Ш. Н.* «Логистика негіздері» пәні бойынша практикалық сабақтарды жүргізу және студенттердің өзіндік жұмыстарын орындауға арналған әдістемелік нұсқау – (5В090900 «Логистика» мамандығының студенттері үшін). − Алматы: ҚазҰТЗУ, 2015. – 15 б.

37. https://kazdata.kz

38. https://stat.gov.kz

39. https://logistics.**kz**

МАЗМҰНЫ

	Кіріспе……	3
1.	ЭКОНОМЕТРИКАҒА КІРІСПЕ.....	4
1.1.	Ықтималдық теориясының және математикалық статистиканың элементтері.....	5
1.2.	Жиынтық сандық сипаттамалары	9
1.3.	Кездейсоқ айнымалының тұрақты және айнымалы құрамдастары.....	12
1.4.	Ковариация, дисперсия және корреляция коэффициенті.....	16
1.5.	Жұптық регрессия.....	21
1.6.	Регрессияның жұптық сызықтық моделі және корреляция.....	25
1.7.	Қалыпты үлестірімді шаманың параметрлерін интервалдық бағалау.....	33
1.8.	Жұптық сызықтық регрессия параметрлерін есептеу.....	36
1.9.	Өзара тәуелділік байланысын зерттеудегі корреляциялық-регрессиялық әдісімен модельдеу	43
1.10.	Сызықтық регрессия коэффициенттерінің статистикалық маңыздылығы.....	49
1.11.	Монте-Карло әдісі бойынша эксперимент.....	52
1.12.	Гаусс-Марковтың 4-шарты.....	57
1.13.	Сызықтық емес экономикалық модельдер.....	65
1.14.	Кобба-Дугластың өндірістік функциясы.....	73
1.15.	Жиынтық регрессия және корреляция.....	76
1.15.1.	Ең кіші квадраттар әдісі (ЕКӘ). ЕКӘ негізінде алынатын бағалардың қасиеттері.....	81
1.15.2.	Факторлардың мәнділігін тексеру және регрессия сапасының көрсеткіштері.....	88
2.	ЛОГИСТИКАЛЫҚ ЖҮЙЕЛЕРДЕГІ БОЛЖАУ ӘДІСТЕРІ.....	99
2.1.	Болжау жасау теориясының негізгі ережелері.....	99
2.2.	Уақыттық қатарлар ……...	104

2.3. Уақыттық қатарлардың компоненттері. Тенденцияның бар болуы туралы гипотезаны тексеру 108
2.4. Динамиканың негізгі көрсеткіштері және уақыттық қатарларды тегістеу үшін жылжымалы орташаларды қолдану..... 112
2.5. Жылжымады орташа көмегімен уақыттық қатарларды тегістеу..... 122
2.6. Өсу қисықтары бар моделдің көмегімен дамуды болжау 132
2.7. Логистиканы болжауда өсу қисықтарының модельдерін қолдану 140
2.8. Логистикалық есептерді болжау үшін экспоненциалды тегістеу және адаптивті әдістерді пайдалану 155
2.9. Жоғары реттегі адаптивті полиномиалды модельдер..... 167
Библиографиялық тізім..... 176

Оқулық басылым

Рауза Төлтайқызы Исмаилова

ЛОГИСТИКАЛЫҚ ЖҮЙЕЛЕРДЕГІ БОЛЖАУ ӘДІСТЕРІ

Оқулық

БДДБ бастығы *Л. Т. Қасжанова*
Редакторы *Т. С. Жақсыбаева*
Компьютерде беттеген *Л. Т. Қасжанова*

Басуға 21.11.2017 ж. қол қойылды
Таралымы 300 дана. Пішімі 60х84 1/16
№ 1 баспаханалық қағаз.
Көлемі 11,3 есепті-баспа табақ. 10,5 шартты баспа табақ.
Тапсырыс № 824. Бағасы келісімді

Қ. И. Сәтбаев атындағы
Қазақ ұлттық техникалық зерттеу университеті басылымы,
Маркетинг және коммуникациялар департаменті
Баспа басқармасы
Алматы қ., Сәтбаев көшесі, 22

ISBN 978-601-323-059-7

