

ПОЗНАВАТЕЛНИ ВЪЗМОЖНОСТИ НА КОЕФИЦИЕНТА НА АВТОДЕТЕРМИНАЦИЯ ПРИ СТАТИСТИЧЕСКИЯ АНАЛИЗ

*Любомир Иванов**



Въведение

Социално-икономическите явления не са постоянни във времето, а се променят непрекъснато, като динамиката им се обуславя както от систематични фактори с определящо влияние, така и от странични причини, имащи случаен характер. Класическата постановка при анализа на динамичните редове разглежда развитието като формирано от четири компонента - тенденция, сезонни колебания, циклични колебания и случайни отклонения. Част от тези компоненти имат систематичен характер и могат да се прогнозират с относително висока точност - тенденцията и сезонните колебания, или с по-ниска точност - цикличните колебания, докато в същото време случайните отклонения не могат да се прогнозират, а се установяват единствено границите, в които варират с определена вероятност. По този начин възможностите за статистическия анализ на динамичните редове и прогнозирането на бъдещото им изменение са директно свързани с наличието на закономерности в развитието, изразени чрез отделните компоненти и съотношенията между тях.

В предходни публикации предложих да се използва коефициентът на

* Доцент д-р, Стопанска академия „Д. А. Ценов“ - Свищов; e-mail: lubomir.ivanov@uni-svishtov.bg.

автодетерминация като обобщен измерител на обективните закономерности в развитието (Иванов, 2010) и изведох статистически тестове за неговата статистическа значимост (Иванов, 2014). Трябва да се посочи, че освен като измерител на закономерностите в динамичните редове коефициентът на автодетерминация носи полезна информация и за отделните компоненти на развитието, за възможностите за прогнозиране и точността на прогнозите, за случайността на динамичните редове. Във връзка с това е разработен настоящият материал, чиято цел е да разкрие основните насоки, в които може да се използва коефициентът на автодетерминация при извършване на емпиричен анализ на развитието на социално-икономическите явления и процеси на основата на динамични редове. Изложението е структурирано в две части. В първата е дадена кратка характеристика на коефициента на автодетерминация, начина за изчисляването му и за определяне на статистическата му значимост. Във втората част са систематизирани петте основни аспекта на познавателното значение на коефициента, илюстрирани с приложение на коефициента на автодетерминация при анализа на конкретни динамични редове.

1. Дефиниране, оценка и статистическа значимост на коефициента на автодетерминация

Коефициентът на автодетерминация (означен като *AUD*) е статистическа относителна величина, която дава обобщена оценка на съществуващите вътрешни закономерности в развитието на социално-икономическите явления и процеси, представени чрез статистически динамични редове. Той е построен по аналогия с коефициента на детерминация, който се използва при регресионния анализ на връзки и зависимости. Между двата коефициента - на детерминация и на автодетерминация, има общи черти, но и определени различия. Общото е, че и двата коефициента се изчисляват във формата на отношение между показатели за разсейването в интересуващите ни признаци - дисперсиите. При това и в двата случая общата дисперсия се разлага на две части, като едната е породена от случайните фактори, докато втората е формирана под влиянието на съществените, закономерни причини. В резултат на този еднотипен начин на конструиране и двата показателя приемат само положителни значения и се движат в границите между нула и единица. Значение „нула“ отразява липсата на закономерности, а при величина „единица“ се наблюдава функционална

зависимост или напълно определени закономерности. Колкото величината на коефициентите е по-близо до нулата, толкова по-голям е дялът на случайността в общото разсейване, а колкото величината им се доближава до единицата, толкова по-голям дял от вариацията или промените в значенията на признака се дължи на действието на обективните закономерности, които в случая с коефициента на автодетерминация приемат формата на систематични причини, определящи основните характеристики на развитието.

Между двата коефициента съществуват и съществени различия. На първо място, при коефициента на определение се използват два или повече признака едновременно, като се дефинират причина и следствие, респективно фактор и резултат. Коефициентът на автодетерминация се използва при отделни, взети сами за себе си явления и процеси, представени само с един динамичен ред от значения на признака. Един и същ признак играе ролята и на фактор, и на резултат, като причината и следствието се дефинират не по отношение на различни свойства, а по отношение на различие във времето - настояще и минало на един и същ признак. На второ място, при коефициента на определение се построява конкретен модел на зависимостта между наблюдаваните признаци и коефициентът е валиден само в рамките на този модел. В същото време коефициентът на автодетерминация не изисква да се построява модел на зависимостта, а той самият дава допълнителна информация за това доколко е удачно да се използва модел за описание на закономерностите в развитието.

Коефициентът на автодетерминация се дефинира като отношение на систематичната дисперсия (или дисперсията, породена от систематичните причини - тенденция, сезонни и циклични фактори) и общата дисперсия на реда:

$$AUD = \frac{\sigma_g^2}{\sigma_y^2},$$

където:

AUD е коефициентът на автодетерминация;

σ_g^2 - систематичната дисперсия;

σ_y^2 - общата дисперсия на динамичния ред.

Оценката на коефициента на автодетерминация може да се извърши по различни начини. Изчислен на базата на първите p коефициенти на автокорелация, той има следния вид за стационарни редове (Иванов, 2010,

с. 13):

$$AUD = \sum_{j=1}^p \varphi_j \rho_j = R' . RR^{-1} . R,$$

където:

ρ_j са коефициентите на автокорелация;

φ_j - частните автокорелационни коефициенти;

$$R = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \dots \\ \rho_p \end{bmatrix} \text{ - векторът стълб с автокорелационните коефициенти;}$$

$$RR = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{p-2} \\ \rho_2 & \rho_1 & 1 & \dots & \rho_{p-3} \\ \dots & \dots & \dots & \dots & \dots \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \dots & 1 \end{bmatrix} \text{ - симетричната квадратна матрица с}$$

автокорелационните коефициенти.

При нестационарни редове коефициентът на автодетерминация се изчислява на основата на:

$$AUD = 1 - \frac{\sigma_{\Delta y, k}^2}{\sigma_y^2} \left[1 - R'_{d, k} . RR_{d, k}^{-1} . R_{d, k} \right],$$

където:

R_d и RR_d са дефинираните по-горе матрици на автокорелационните коефициенти с тази разлика, че са изчислени на основата на k -тите разлики, а не на нестационарния ред;

$\sigma_{\Delta y, k}^2$ - дисперсията на реда на k -тите разлики, който е стационарен (Иванов, 2010, с. 15).

Проверката на статистическата значимост на коефициента на автодетерминация се основава на тестовата характеристика (Иванов, 2014, с. 78):

$$Q_{(p)} = n . AUD,$$

където:

n е дължината на динамичния ред;

p - броят на използваните при оценката автокорелационни коефициенти.

Тестовата величина $Q_{(p)}$ следва χ^2 - разпределение с p степени на свобода. Коефициентът се приема за статистически значим, когато емпиричната стойност на тестовата характеристика надвишава табличната при определен риск от грешка. Когато емпиричната стойност е по-малка или равна на теоретичната, то коефициентът не може да се отличи от нулата. В този случай коефициентът не е отчел наличие на съществени закономерности в развитието на явленията.

2. Познавателна същност на коефициента на автодетерминация

2.1. Мярка на закономерностите в развитието

Първият аспект от познавателната същност на коефициента на автодетерминация е като **мярка на обективните закономерности в развитието**. По начина на своето конструиране коефициентът приема стойности в интервала $[0; 1]$, като значение на коефициента $AUD = 0$ показва липса на закономерности в развитието (напълно случаен процес)¹, а значение $AUD = 1$ показва пълна обусловеност на изучаваното явление от неговото минало (детерминизъм). Колкото величината на коефициента е по-близо до единица, толкова по-силно изразени са вътрешните закономерности в развитието. Колкото величината е по-близо до нулата, толкова по-слаби са наличните закономерности.

Коефициентът на автодетерминация показва каква част от вариацията се дължи на съществуващите обективни закономерности в развитието на явленията и каква - на случайни причини. Когато значението му надвишава **0.5**, може да се каже, че преобладава влиянието на систематичните фактори, а в случай че величината на коефициента е по-малка от **0.5**, преобладава влиянието на случайността, на непредсказуемата неопределеност. Подобен извод е валиден по отношение на промените в реда като цяло, а не по отношение на отделните индивидуални изменения за всеки отделен период (или момент) спрямо предходния.

¹ Необходимо е да поясним, че един напълно случаен процес („бял шум“), разгледан сам по себе си като последователност от реализации на нормално разпределена случайна величина (респективно на независими и идентични случайни величини), на практика по нищо не се различава от един ред на разпределение. Той притежава определени закономерности, свързани със закона на разпределението на случайната величина - функцията на плътността на вероятностите и нейните параметри. Тези закономерности обаче не са свързани с динамиката на явлениято, при тях липсва обусловеността на настоящето от миналото (респективно на бъдещето от настоящето), поради което по-нататък в изложението ще се абстрахираме от тях.

Допълнителна информация за компонентите на развитието може да се получи при сравнение на коефициентите на автодетерминация, изчислени от началните и от трансформирани данни в случаите, когато се регистрира присъствие на нестационарност. Високите, статистически значими стойности на коефициента за трансформирани данни показват наличие на значими циклични и/или сезонни компоненти. Комбинацията от близък до нулата коефициент при трансформирани данни и статистически значим коефициент за началните данни показва наличие само на тенденция на развитие (табл. 1). Значими стойности на коефициенти при началните и при трансформирани данни са индикация за наличие както на тенденция, така и на циклични и/или сезонни компоненти. По този начин може да се даде отговор на въпроса дали закономерностите се формират само от тенденцията, или и от сезонни и циклични компоненти, както и да се оцени сравнителната сила на отделните компоненти на развитието.

1. Тълкуване на възможните комбинации на коефициента на автодетерминация при началните и трансформирани данни

Коефициент на автодетерминация в началните данни	Коефициент на автодетерминация в трансформирани данни	
	значим	незначим
Значим, съществено надвишаващ коефициента в трансформирани данни	Тенденция и сезонност и/или цикличност	Само тенденция
Значим, несъществено надвишаващ коефициента в трансформирани данни	Само сезонност и/или цикличност	х
Незначим	х	Няма закономерности

В табл. 2 са представени стойностите на коефициента на автодетерминация за динамичните редове на три демографски показателя: „коефициент на раждаемост“ n (фиг. 1 от приложението), „брой на живородените момчета на 1 000 родени момичета“ $N_{m/f}$ (фиг. 2 от приложението) и „брой на починалите мъже на 1 000 починали жени“ $M_{m/f}$ (фиг. 3 от приложението) за периода 1930 - 2013 година. Използвани са

демографски показатели, тъй като, от една страна, за тях има на разположение достатъчно дълги динамични редове, за които са изпълнени изискванията за съпоставимост, а от друга - в тях присъстват систематични закономерности, което ги прави особено подходящи за илюстриране на възможностите на коефициента на автодетерминация.

Величините са изчислени след предварителен анализ за наличие на тенденция, при което за два от редовете е установено, че са нестационарни и е използвана съответната модифицирана формула. На основата на анализ на статистическата значимост на автокорелационните и частните автокорелационни коефициенти на редовете (или на стационарните им части след трансформацията) е определен броят на включените коефициенти - съответно до порядък **17.8** и **2**.

2. Коефициенти на автодетерминация за някои демографски показатели за периода 1930 - 2013 година²

Показатели	Коефициент на автодетерминация
Коефициент на раждаемост n - ‰	0.988
Живородени момчета на 1 000 родени момичета $N_{m/f}$ - бр.	0.239
Починали мъже на 1 000 починали жени $M_{m/f}$ - бр.	0.744

Най-силни вътрешни закономерности се наблюдават в реда на коефициента на раждаемост n - **98.8%** от измененията се обуславят от систематичните причини в динамиката, а едва **1.2%** от промените се дължат на случайните фактори. Систематичните компоненти имат определящо влияние при формирането на реда, докато влиянието на случайността е пренебрежимо малко. В същото време в реда на живородените момчета на 1 000 родени момичета $N_{m/f}$ закономерностите в развитието обуславят по-малко от една четвърт от промените - само **23.9%** от вариацията се дължат на систематичните причини. В този ред случайността има определящо влияние и доминира над постоянно действащите фактори. За третия динамичен ред - на починалите мъже на 1 000 починали жени $M_{m/f}$, се установява, че закономерните причини

² Повече подробности относно изчисляването на коефициентите на автодетерминация за посочените динамични редове могат да се проследят в: Иванов, 2014, с. 90 - 94.

обуславят почти три четвърти от промените - **74.4%**. Тяхното влияние е по-силно отколкото влиянието на случайните фактори, въпреки че значението на последните също не е за пренебрегване - те формират **25.6%** от вариацията на реда.

Следователно на базата на величините на коефициентите на автодетерминация може да се установи, че най-силни закономерности се наблюдават в развитието на коефициента на раждаемост n , по-слаба, но също значителна е силата на закономерностите в динамиката на починалите мъже на 1 000 починали жени $M_{m/f}$, а най-слаби закономерности има в динамиката на показателя „живородени момчета на 1 000 родени момичета“ $N_{m/f}$.

3. Коефициенти на автодетерминация за първите разлики на нестационарните демографски показатели за периода 1930 - 2013 година³

Показатели	Коефициент на автодетерминация
Коефициент на раждаемост n - ‰	0.510
Починали мъже на 1 000 починали жени $M_{m/f}$ - бр.	0.079

В табл. 3 са посочени коефициентите на автодетерминация на трансформираните редове. При сравнение със значенията на коефициентите за началните данни (табл. 2) може да се установи, че за показателя „починали мъже на 1 000 починали жени“ $M_{m/f}$ закономерностите се формират единствено от тенденцията - незначим коефициент на трансформираните данни и значим коефициент за началните данни. В същото време при показателя „коефициент на раждаемост“ n закономерностите се формират както от циклични колебания, така и от тенденцията - значим коефициент при трансформираните данни и значим и със значително по-голяма стойност коефициент при началните данни.

2.2. Мярка на потенциалната възможност за построяване на линеен иконометричен модел, описващ закономерностите в развитието

Всеки модел представлява опростена картина на действителността. Иконометричните модели представят съществуващите закономерности и следователно могат да пресъздадат реда само до рамките на тези

³ Повече подробности относно изчисляването на коефициентите на автодетерминация за посочените динамични редове може да се проследят в: Иванов, 2014, с. 90 - 94.

закономерности. Когато в динамичния ред не се наблюдават съществени, вътрешно присъщи закономерности в развитието, построяването на модел става безпредметно, тъй като той не би могъл да представи нещо, което не съществува. Такъв модел не може да бъде адекватен и съответно няма да бъде статистически значим. И обратното, когато в реда съществуват силно изразени закономерности, когато делът на систематичната вариация е висок, тогава могат да бъдат построени един или повече иконометрични модели, които да пресъздадат наличните вътрешни зависимости в реда. Дали всеки от тях ще е адекватен, не може да се каже предварително, но високата, статистически значима величина на коефициента на автодетерминация означава, че поне един адекватен модел може да бъде построен. Следователно вторият аспект от познавателната същност на коефициента на автодетерминация е **мярка на потенциалната възможност за построяване на линеен иконометричен модел, описващ закономерностите в развитието**. Като представя в количествена форма отношението между систематичната и общата дисперсия на началните и на трансформираните данни, коефициентът на автодетерминация дава представа доколко наличните закономерности се формират само от тенденцията или има съществени сезонни и/или циклични компоненти. По този начин изследователят може да направи извод дали за представяне на съществуващите в реда закономерности е необходимо да се построи модел само на тенденцията на развитие, дали е достатъчен само стационарен модел на цикличните и/или сезонните компоненти, или е нужно да се построи обобщен модел, който да отразява както наличната тенденция, така и присъствието на сезонност и/или цикличност.

Анализът на значението на коефициента на детерминация позволява да се направи извод за степента, до която един модел може потенциално да пресъздаде реда, респективно - за силата на случайните колебания, които не могат да се моделират на базата на трендови модели или на модели на сезонността и/или цикличността. В този аспект коефициентът дава потенциала - границата, до която може да достигне коефициентът на детерминация при построяването на единичен линеен модел. Ако има необходимост да се построи модел, който да пресъздава по-добре реда, да дава по-висока стойност на обяснената вариация, това не може да стане на базата на информацията в реда сама по себе си. За да се построи по-добър модел, е нужна повече информация,

следователно, ако целта е да се построи модел, който да обяснява по-голяма част от вариацията на реда, отколкото е значението на коефициента на автодетерминация, то трябва да се използват **допълнителни данни** - други динамични редове, които представят явления или процеси, обвързани по някакъв начин с изучаваното явление.

Всички тези заключения са валидни само за линеен модел. Потенциалът за моделиране е потенциал за построяване на **линеен модел** на закономерностите. Това не изключва възможността модел с някаква друга функционална форма (нелинейна) да има по-висок потенциал за обяснението на промените в реда. Ако данните се трансформират по подходящ начин (логаритмична, реципрочна, логистична и друга трансформация), те могат да се линеаризират и да се изчисли коефициентът на автодетерминация от трансформираните данни. Той обаче ще има различно тълкуване, тъй като вариацията на реда след трансформацията е различна и следователно пропорцията от нея, измерена от коефициента, е при различна основа. Коефициентът може да се използва за сравнение при различни трансформирани редове само в случай, че при всички тях се използва една и съща трансформация.

Анализът на получените стойности на коефициента на детерминация в табл. 2 и 3 показва, че коефициентът на раждаемост n има най-висок потенциал за построяване на линеен модел - той би обяснил **98.8%** от вариацията в реда. Тъй като в реда се съдържа значим цикличен компонент в стационарната част (**0.510**), то в модела трябва да се включи както компонент на тенденцията, така и компонент, отразяващ цикличните колебания. Доколкото единичен модел би обяснил почти **99%** от вариацията, а за случайните колебания остава около **1%**, то включването на допълнителни данни - други динамични редове, не би оказало съществено влияние за подобряване на модела и не е оправдано от гледна точка на намаляването на степените на свобода.

В динамичния ред на показателя „починали мъже на 1 000 починали жени“ $M_{m/f}$ може да се моделират до **74.4%** от промените. Тъй като стационарната част не притежава статистически значим коефициент на автодетерминация (**0.079**), то моделът трябва да съдържа само компонент на тенденцията. Случайните колебания обхващат **25.6%**, което дава сериозни възможности за подобряване на един линеен модел с включване на информация

от други динамични редове.

Третият ред на показателя „живородени момчета на 1 000 родени момчета“ N_{mf} е стационарен и не съдържа значим циклически компонент ($AUD = 0.239$). В най-добрия случай могат да се пресъздадат до **23.9%** от промените в реда. За да се постигне по-добър модел, е необходимо да се използва допълнителна информация от други динамични редове. На практика без подобна допълнителна информация няма никаква гаранция, че единичен модел би бил въобще адекватен, тъй като коефициентът на автодетерминация е на границата на статистическата значимост.

2.3. Мярка на прогнозируемостта на динамичните редове

На трето място, коефициентът на автодетерминация може да се използва за оценка на прогностичните възможности на динамичния ред. Поради спецификата на динамичните редове, и по-специално на качеството им „подреденост“, отразяващо хронологичната последователност на отделните им елементи, наличието на вътрешни закономерности в развитието се трансформира в наличието на обективно обусловена зависимост на настоящето от миналото. Всеки от членовете на реда е свързан с предхождащите го членове и в същото време предопределя в известни граници следващите го членове. Като измерва силата на съществуващите закономерности, коефициентът на автодетерминация в същото време дава оценка и на зависимостта на бъдещето от настоящето и миналото, тъй като отразява каква част от промените в реда се обуславят от неговото минало. Когато величината на показателя е близка до нулата, редът не може да се прогнозира точно, тъй като случайността доминира. В този случай опитите да се прогнозира бъдещото развитие са безпредметни, тъй като в реда не се съдържа достатъчно информация за неговото бъдеще. Когато стойността се доближава до единица, закономерностите в развитието на реда са силни и прогнозирането е по-точно. Наличието на статистически значима стойност на коефициента означава, че в реда се съдържа достатъчно информация, за да се прогнозира бъдещото му развитие, при условие че не настъпят радикални промени в характера на съществуващите обективни закономерности. Следователно третият аспект от познавателната същност на коефициента на автодетерминация е **мярка на прогнозируемостта на динамичните редове.**

Колкото по-висока стойност има коефициентът на автодетерминация, толкова по-малка е потенциалната величина на стандартната грешка на оптималния модел, следователно толкова по-малка ще е и относителната прогностична грешка. Трябва да се има предвид, че абсолютната грешка на прогнозата зависи и от величината на дисперсията на реда. Само когато се сравняват два реда с равни или приблизително равни дисперсии съотношението между относителните и абсолютните грешки на прогнозата ще е едно и също - по-малка е потенциалната грешка при прогнозирането на този ред, за който коефициентът на автодетерминация е по-голям.

Сравнението на коефициентите на автодетерминация в табл. 1 показва, че с най-висока относителна точност може да се прогнозира редът на показателя „коефициент на раждаемост“ n , тъй като потенциалната грешка на модела ще се определя от остатъчните 1.2% от дисперсията на реда. В същото време при показателя „починали мъже на 1 000 починали жени“ $M_{m/f}$ относителната грешка ще е по-голяма, тъй като се базира на остатъчните 25.6% от дисперсията на реда. Най-ниска степен на прогнозируемост има редът на показателя „живородени момчета на 1 000 родени момичета“ $N_{m/f}$. Потенциалната относителна грешка ще се изчислява на базата на 76.1% от дисперсията на реда, и то при условие, че се построи модел, който да обхване напълно тези 23.9%, които се дължат на систематичните причини в развитието. Сравнението между показателите и характеристиките на тяхното развитие във времето, изразени в коефициента на автодетерминация, води до извода, че прогнозирането ще е ефективно при реда на коефициента на раждаемост n , с приемлива точност при показателя „починали мъже на 1 000 починали жени“ $M_{m/f}$ и на практика ще е безпредметно при реда на показателя „живородени момчета на 1 000 родени момичета“ $N_{m/f}$.

2.4. Критерий за случайност на динамичните редове

На четвърто място, коефициентът на автодетерминация дава възможност да се прецени доколко динамичният ред е случаен. Когато един ред е случаен, той не съдържа систематичен компонент. В този случай вариацията на реда не се разпада на две части, а се трансформира единствено във вариация на случайния компонент. Коефициентът на автодетерминация ще приеме стойност, равна на нула. На практика е възможно показателят да има величина, която да

не е нула, но да не се отличава съществено от нулата при определен риск от грешка. Тогава коефициентът на автодетерминация не е статистически значим. Липсата на статистическа значимост е равнозначна на случайност на динамичния ред, следователно проверката за статистическа значимост на коефициента представлява алтернатива на вече използваните в практиката тестове за случайност на реда. По този начин четвъртият аспект на познавателната същност на коефициента на автодетерминация е **критерий за случайност на динамичните редове**.

Трябва да отбележим, че проверката за случайност на реда не е идентична с проверката за стационарност на реда, тъй като дихотомията „стационарност - нестационарност“ се извежда само по отношение на компонента на тенденцията в реда. Стационарните редове също могат да притежават определени закономерности, изразени в наличието на сезонен или цикличен компонент. Проверката за случайност е по същество проверка за наличието на всички систематични компоненти на реда, изразяващи присъщите на развитието закономерности, включително и на тенденцията. Следователно когато величината на коефициента на автодетерминация не е статистически значима, това означава, че редът е стационарен, докато обратното не е валидно - наличието на стационарност не означава непременно незначима стойност на коефициента на автодетерминация.

Когато редът е случаен, той става сходен с вариационните редове на разпределение и като такъв се характеризира със средно значение и граници, в които варира. Движението във времето не променя нито центъра на разпределението, нито величината на разсейването в реда, доколкото те са инвариантни по отношение на времето.

Проверката за случайност на динамичните редове се базира на тестовата характеристика:

$$Q_{(p)} = n \cdot AUD,$$

следваща χ^2 - разпределение с p степени на свобода.

В табл. 4 са представени тестовите характеристики за трите анализирани демографски показателя. Проверката на хипотезата за случайност при 5% риск от грешка води до заключението, че и трите реда не са напълно случайни, а съдържат някакви закономерности в развитието си. Те не могат да се третират

като идентични с редовете на разпределение и следователно могат да се анализират по отношение на построяването на модели на динамика и прогнозиране на бъдещото им развитие.

4. Статистическа значимост на коефициентите на автодетерминация за някои демографски показатели за периода 1930 - 2013 година

Показатели	Емпирична характеристика	Теоретична характеристика (5% риск от грешка)
Коефициент на раждаемост n - ‰	82.984	27.587
Живородени момчета на 1 000 родени момичета $N_{m/f}$ - бр.	20.090	15.507
Починали мъже на 1 000 починали жени $M_{m/f}$ - бр.	62.524	5.991

2.5. Основа за сравнение при определяне на адекватността на модела

На пето място, коефициентът на автодетерминация представлява база за сравнение на различни динамични модели. При иконометричното моделиране на закономерностите в развитието, в т.ч. и на тенденцията на развитие, често се използват различни конкуриращи се модели. Техните характеристики се сравняват и се подбира този от тях, който пресъздава най-точно динамичния ред. Като един от критериите за избор на модел се използва коефициентът на определение (Величкова, 1981, с. 112 - 113). Трябва да се има предвид, че различните модели дават различни значения на коефициента на определение. В същото време коефициентът на автодетерминация измерва обобщено всички закономерности на развитието, следователно представлява горна граница, до която биха могли да достигнат значенията на коефициента на определение, ако съответният модел пресъздава точно обективните закономерности в развитието. Колкото по-близо е значението на коефициента на определение до величината на коефициента на автодетерминация, толкова по-адекватен е моделът, респективно - колкото по-голяма е разликата между двете величини, толкова повече възможности има за подобряване на модела. Следователно величината на коефициента на определение не бива да се оценява абсолютно, сама по себе си, и да се твърди, че щом тя е ниска, то моделът може да се подобри. Моделът може да се подобри само ако коефициентът на определение се различава

съществено от коефициента на автодетерминация. Следователно петият аспект на познавателната същност на коефициента на автодетерминация е като **основа за сравнение при определяне на адекватността на модела.**

В този аспект коефициентът на автодетерминация е полезно допълнение при диагностичния анализ на адекватността на модела, който трябва да се използва съвместно с проверката за адекватност и величината на коефициента на детерминация на регресионния модел. Той позволява да се даде отговор на редица въпроси, свързани с адекватността, например има ли възможност за подобрене на модела, има ли възможност да се достигне до адекватен модел, ако диагностицираният не е адекватен и т.н.

В табл. 5 са посочени коефициентите на детерминация на линейните динамични модели на трите анализирани демографски показателя. Моделите са от смесен тип - полином от втора степен и авторегресионен компонент от първи порядък. Моделът е построен идентично и за трите показателя, за да е коректно сравнението на получените резултати. Спецификацията е съобразена, от една страна, с индикации за наличие на тенденция, а от друга - с необходимостта да се елиминира автокорелацията в остатъчните елементи, тъй като при наличие на автокорелация оценките на дисперсията, а следователно и на коефициента на детерминация на модела, стават изместени (Hamilton, 1994, с. 208 - 214).

Решаването на моделите дава следните резултати:

Коефициент на раждаемост

$$n_t = 4,9137 - 0,0621 \cdot t + 0,0003 \cdot t^2 + 0,8121 \cdot n_{t-1};$$

Живородени момчета на 1 000 родени момичета

$$N_{m/f,t} = 932,1009 - 0,1806 \cdot t + 0,0011 \cdot t^2 + 0,1266 \cdot N_{m/f,t-1};$$

Починали мъже на 1 000 починали жени

$$M_{m/f,t} = 203,6822 + 0,7845 \cdot t - 0,0085 \cdot t^2 + 0,8120 \cdot M_{m/f,t-1}.$$

5. Коефициенти на определение на моделите на демографските показатели за периода 1930 - 2013 година

Показатели	Коефициент на определение ⁴	Коефициент на автодетерминация
Коефициент на раждаемост n - ‰	0.977	0.988
Живородени момчета на 1 000 родени момичета $N_{m/f}$ - бр.	0.095	0.239
Починали мъже на 1 000 починали жени $M_{m/f}$ - бр.	0.737	0.744

Разгледани сами по себе си, коефициентите на определение биха показали, че за коефициента на раждаемост е построен адекватен модел, докато при останалите два показателя има възможности за подобрене, особено при реда на живородените момчета на 1 000 родени момичета. Когато сравним получените величини на коефициента на определение със значенията на коефициента на автодетерминация, са оказва, че както при реда на коефициента на раждаемост, така и при реда на починалите мъже на 1 000 починали жени моделите са достигнали изключително близо до възможната граница. Дори и да се опитваме да построим по-добър модел за третия показател, това не е възможно при наличните данни и трансформации. Просто в реда има определена закономерност и посоченият модел я изчерпва на практика напълно. Известна възможност за подобрене има само в модела на втория показател „живородени момчета на 1 000 родени момичета“. Дори и в този случай обаче сравнението със значението на коефициента на автодетерминация е полезно, тъй като показва, че подобряването на модела не може да достигне значения на определението от порядъка на **0.977** или **0.737** както при другите два модела. Най-доброто, което може да се очаква, е модел с коефициент на определение от порядъка на **0.200 - 0.230**.

⁴ Използвани са коригираните коефициенти на определение, които са по-точни измерители на адекватността на модела (Maddala, 1988, с. 126).

Заклучение

Основната теза, застъпена в настоящата статия, е, че коефициентът на автодетерминация дава съществени възможности за задълбочаване и разширяване на статистическия анализ на динамика. Във връзка с това могат да се направят следните **изводи от теоретико-методологично естество**, свързани с използването на коефициента в следните направления:

Първо, той позволява да се разграничават динамичните редове, съдържащи систематични компоненти, от напълно случайните редове и следователно оценява необходимостта от построяване на модел за пресъздаване на закономерностите.

Второ, коефициентът дава полезна информация за характеристиките на динамичните редове, като посочва компонентите на развитието, които трябва да се включат в модела, за да се постигне адекватно изглаждане на реда.

Трето, коефициентът на автодетерминация показва дали в реда се съдържа достатъчно информация, за да се построи адекватен модел на развитието, или информацията е недостатъчна и в този случай трябва да се търси допълнителна информация под формата на други динамични редове.

Четвърто, коефициентът на автодетерминация дава оценка на потенциала на динамичните редове по отношение на прогнозирането на тяхното бъдещо развитие.

Пето, при построяване на модел сравнението на коефициента на определение с коефициента на автодетерминация позволява да се направи точна преценка не само дали построеният модел е адекватен, а и доколко той може да бъде подобрен чрез промяна на спецификацията му.

Извършеният с помощта на коефициента на автодетерминация емпиричен анализ на закономерностите в развитието на демографските показатели „коефициент на раждаемост“, „живородени момчета на 1 000 живородени момичета“ и „починали мъже на 1 000 починали жени“ в България за периода 1930 - 2013 г. дава възможност да се направят следните **изводи с приложен характер**:

Първо, динамичният ред на показателя „коефициент на раждаемост“ съдържа силни закономерности в развитието си, изразени в наличието на тенденция и цикличен компонент. Те позволяват да се построи адекватен

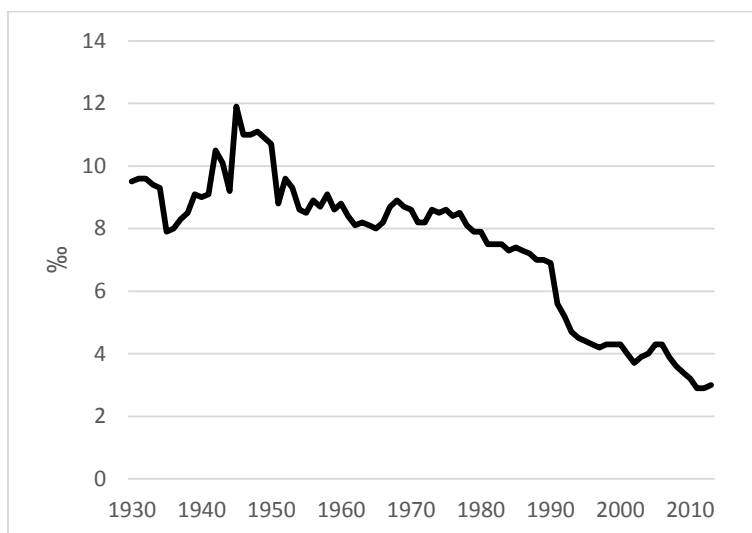
динамичен модел и да се прогнозира бъдещото развитие с висока точност.

Второ, динамичният ред на показателя „живородени момчета на 1 000 родени момичета“ не е напълно случаен, но е стационарен, като наличните закономерности в развитието му са слаби. Построяването на адекватен модел на динамиката му е затруднено и за да се реализира това, както и за да се прогнозира развитието му, е необходимо да се събере допълнителна информация.

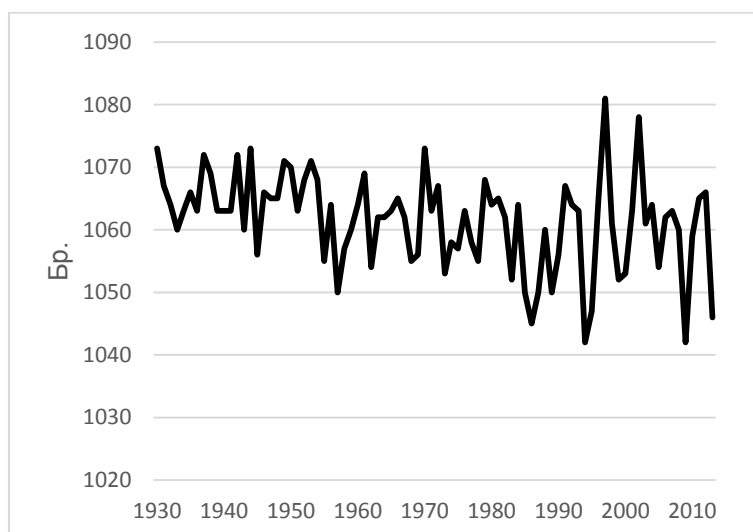
Трето, динамичният ред на показателя „починали мъже на 1 000 починали жени“ съдържа определени закономерности в динамиката си, което позволява да се построи адекватен линеен модел и да се прогнозира бъдещото му развитие. Съществуват възможности за подобрене на подобни модели с включването на допълнителна информация.

Приложение

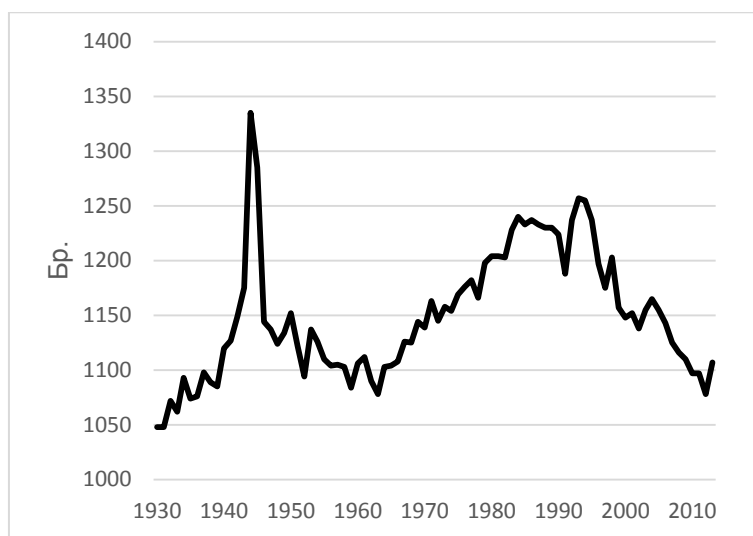
Фиг. 1. Коэффициент на раждаемост в България, 1930 - 2013 година



Фиг. 2. Живородени момчета на 1 000 живородени момичета в България, 1930 - 2013 година



Фиг. 3. Починали мъже на 1 000 починали жени в България, 1930 - 2013 година



ЦИТИРАНА ЛИТЕРАТУРА:

Величкова, Н. (1981). Статистически методи за изучаване и прогнозиране на развитието на социално-икономически явления, София, „Наука и изкуство“.

Иванов, Л. (2010). Относно познавателната същност на автокорелацията в динамичните редове, Статистика, кн. 3 - 4, с. 6 - 28.

Иванов, Л. (2014). Анализ на статистическата значимост на коефициента на автодетерминация, Икономика 21, кн. 2, с. 72 - 101.

Национален статистически институт (2015). Статистически годишник 2014, София.

Box, G. E. P., G. M. Jenkins, G. C. Reinsel (1994). Time Series Analysis: Forecasting and Control. 3rd Edition. New Jersey: Prentice Hall.

Hamilton, D. J. (1994). Time Series Analysis. Princeton: Princeton University Press.

Maddala, G. S. (1988). Introduction to Econometrics. New York: Macmillan.