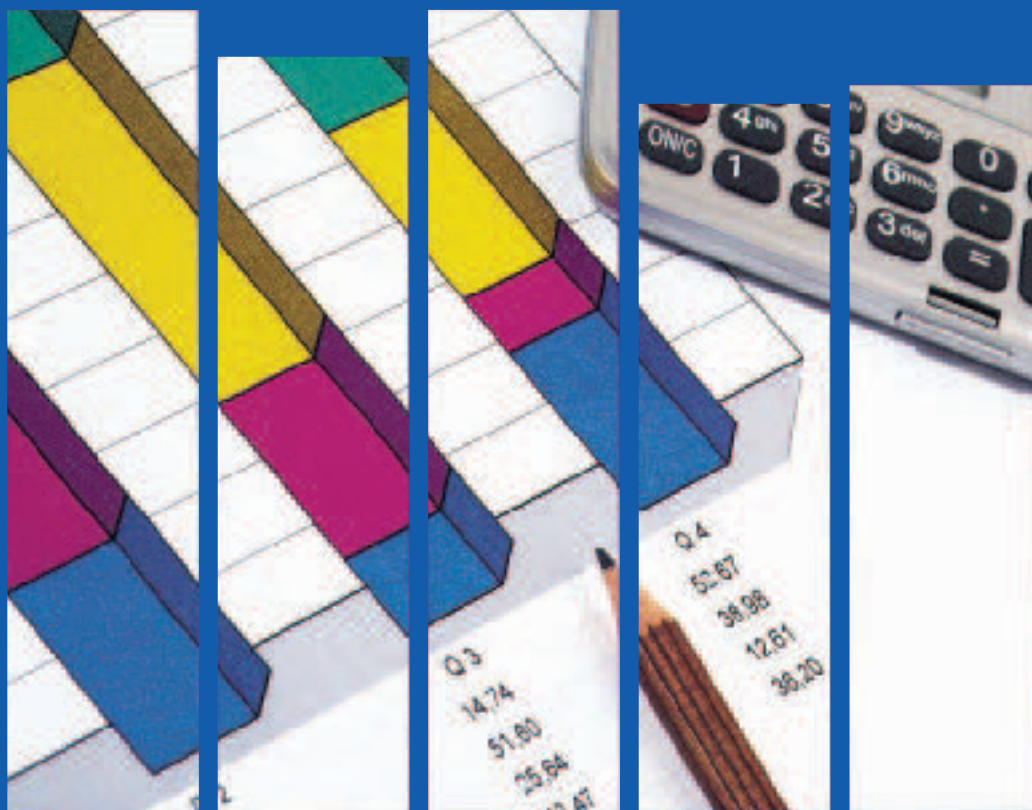
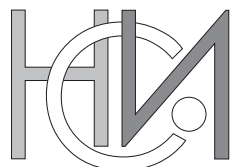


СТАТИСТИКА STATISTICS

1/2018



РЕПУБЛИКА БЪЛГАРИЯ
REPUBLIC OF BULGARIA



НАЦИОНАЛЕН СТАТИСТИЧЕСКИ ИНСТИТУТ
NATIONAL STATISTICAL INSTITUTE

СТАТИСТИКА

STATISTICS

1/2018

СОФИЯ, 2018
SOFIA, 2018

Decorative wavy lines at the bottom of the page, consisting of several overlapping, curved black lines that create a sense of movement and depth.

РЕДАКЦИОННА КОЛЕГИЯ

Главен редактор: д-р Богдан Богданов

Заместник главен редактор: проф. д-р Васил Цанов

Членове:

Проф. д-р Йордан Христосков, проф. д.с.н. Светлана Съйкова, проф. д-р Поля Ангелова,
проф. д-р Димитър Аркадиев, доц. д-р Калоян Харалампиев, доц. д-р Любомир Иванов,
доц. д-р Екатерина Тошева, доц. д-р Александър Цветков, д-р Любен Томев,
д-р Елка Атанасова, д-р Александър Найденов, Стефан Цонев,
Деян Славов, Георги Ангелов

Преводач на руски език: Жаклина Цветкова

Преводач на английски език: Камен Караджов

Отговорен редактор: Лидия Александрова

Стилов редактор: Мила Трифонова

Адрес на редакцията:

София, 1038, ул. „П. Волов” № 2, ет. V

e-mail: bbogdanov@nsi.bg

EDITORIAL TEAM

Chief Editor: Dr. Bogdan Bogdanov

Deputy editor: Prof. Dr. Vasil Tsanov

Members:

Prof. Dr. Jordan Hristoskov, Prof. Dr. Sc. Svetlana Saykova,

Prof. Dr. Polyana Angelova, Prof. Dr. Dimitar Arkadiev,

Assoc. Prof. Kaloyan Haralampiev, Assoc. Prof. Lyubomir Ivanov, Assoc. Prof. Catherine Tosheva,

Assoc. Prof. Alexander Tsvetkov, Dr. Lyuben Tomov, Dr. Elka Atanasova, Dr. Alexander Naidenov,

Stefan Tsonev, Deyan Slavov, Georgi Angelov

Russian Translation: Jaklina Tzvetkova

English Translation: Kamen Karadjov

Editor: Lidia Aleksandrova

Style editor: Mila Trifonova

Editorial address:

2, P. Volov St., Sofia 1038, Bulgaria

Vth floor

e-mail: bbogdanov@nsi.bg

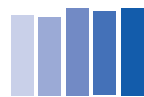


СЪДЪРЖАНИЕ

	Стр.
ТЕОРИЯ И МЕТОДОЛОГИЯ НА СТАТИСТИЧЕСКИТЕ ИЗУЧАВАНИЯ	
Емил Христов	9
Галя Статева	55
СТАТИСТИЧЕСКИ ИЗСЛЕДВАНИЯ И АНАЛИЗИ	
Димитър Радилов	95
Тихомир Върбанов	129
БЪЛГАРСКАТА СТАТИСТИКА В ТРИ СТОЛЕТИЯ	
Тодор Къналиев	155
ИНФОРМАЦИИ, РЕЦЕНЗИИ, КОНСУЛТАЦИИ	
Палмира Фарах	171

СОДЕРЖАНИЕ

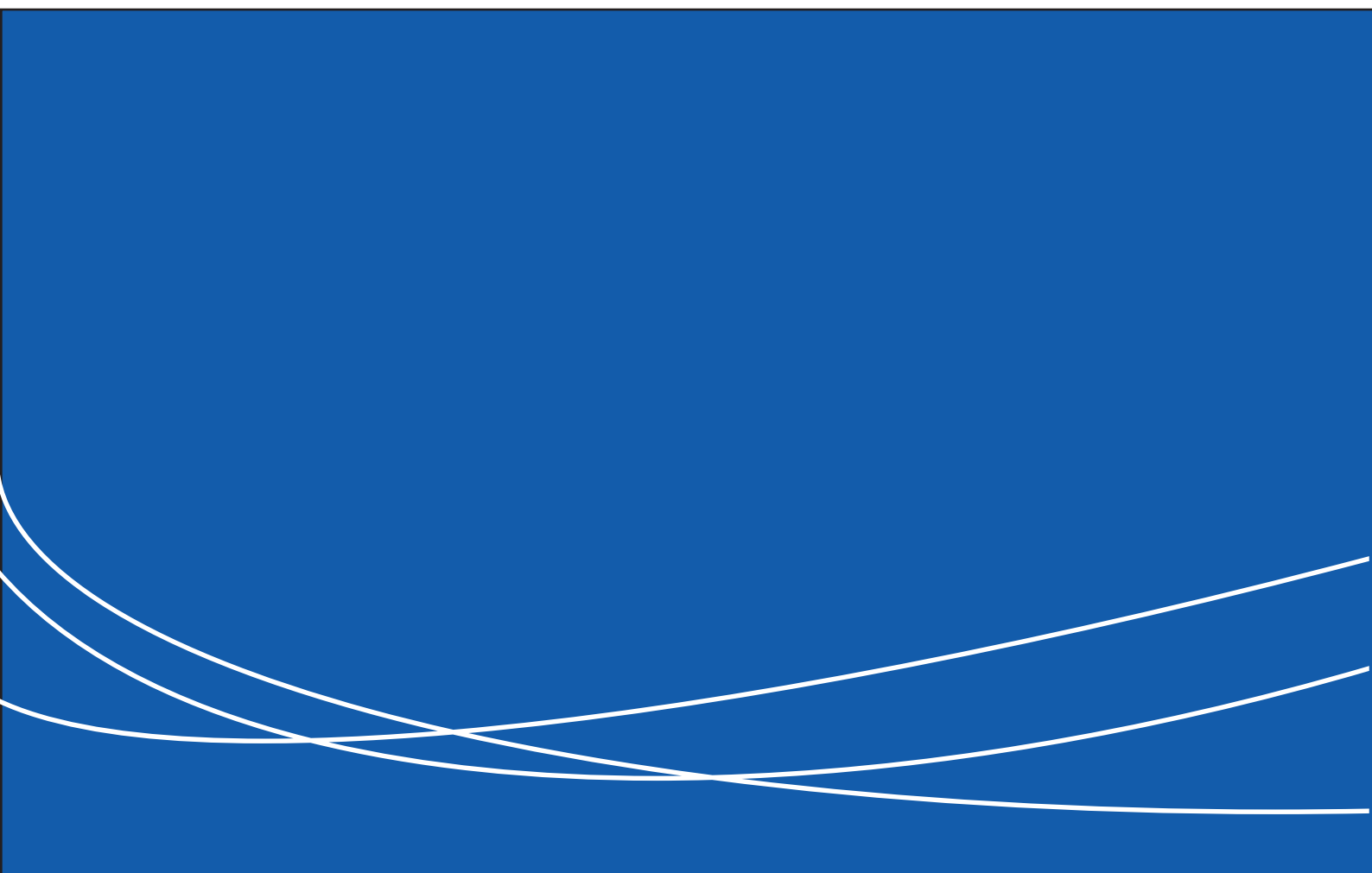
		Стр.
ТЕОРИЯ И МЕТОДОЛОГИЯ СТАТИСТИЧЕСКИХ ИССЛЕДОВАНИЙ		
Емил Христов	Аддитивный факторный анализ средних цен на товары в однородных совокупностях при использовании дискретной нечетной функции математической сигмы	9
Галя Статева	Прикладные аспекты больших данных (Big Data) в официальной статистике	55
СТАТИСТИЧЕСКИЕ ИССЛЕДОВАНИЯ И АНАЛИЗЫ		
Димитр Радилов	Академичные научные позиции в статистической науке и о наступивших переменах в профессии „статистик“ в современном мире	95
Тихомир Вырбанов	Кластеризация европейских стран по степени конкурентоспособности в сфере туризма	129
ПУТЬ БОЛГАРСКОЙ СТАТИСТИКИ ДЛИННОЮ В ТРИ СТОЛЕТИЯ		
Тодор Каналиев	Профессор Венец Цонев - выдающийся творческий путь в теории и практике исследований статистических совокупностей	155
ИНФОРМАЦИИ, РЕЦЕНЗИИ, КОНСУЛЬТАЦИИ		
Палмира Фарах	„Статистика цен на жилье“ - изменения, наступившие в результате внедрения единой европейской рамки. Темы и направления развития, которые обсуждались на международной рабочей встрече в 2016 году (Лиссабон, Португалия, 28-29 апреля 2016 года)	171



CONTENTS

	Page
THEORY AND METHODOLOGY OF THE STATISTICAL SURVEYS	
Emil Hristov	9
Additive factorial analysis of the average prices of homogeneous sets of goods with the discrete non - odd function of the mathematical signum	
Galya Stateva	55
Applicable Big Data aspects in official statistics	
STATISTICAL SURVEYS AND ANALYSIS	
Dimitar Radilov	95
Academic Scientific Positions on Statistical Science and about changes in the profession of statisticians in the contemporary world	
Tihomir Varbanov	129
Clusterization of the European countries based on the level of tourism competitiveness	
BULGARIAN STATISTICS IN THREE CENTURIES	
Todor Kanaliev	155
Prof. Venets Tsonev - a Remarkable Creative Path in the Theory and Practice of Aggregate research	
INFORMATION, REVIEWS, CONSULTATIONS	
Palmira Farah	171
'Housing price statistics' - changes as a result of the implementation of a single European framework. Themes and guidelines for development discussed on the international working meeting in 2016 (Lisbon, Portugal, 28-29 April 2016)	

**ТЕОРИЯ И МЕТОДОЛОГИЯ НА
СТАТИСТИЧЕСКИТЕ ИЗУЧАВАНИЯ**



АДИТИВЕН ФАКТОРЕН АНАЛИЗ НА СРЕДНИТЕ ЦЕНИ ОТ ЕДНОРОДНИ СЪВКУПНОСТИ НА СТОКИ С ДИСКРЕТНАТА НЕЧЕТНА ФУНКЦИЯ НА МАТЕМАТИЧЕСКИЯ СИГНУМ

*Емил Христов**

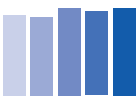


Въведение

В статията е представена **нова методика** за дискретен адитивен факторен анализ на средните цени на стоките в икономическата статистика. Такива цени има обаче само при еднородните статистически съвкупности на стоките. Статистическата съвкупност е основна категория в теорията на статистиката както множеството в теорията на множествата, но за разлика от множествата статистическите съвкупности в икономиката се разделят на **еднородни и разнородни съвкупности**. Тъй като те все още не са точно дефинирани в учебната и научната литература, се използват определенията на автора за еднородна и разнородна съвкупност (Христов, 2016а). Според тези определения под **еднородна съвкупност** се разбира крайно множество на **взаимозаменяеми стоки от един и същ вид** за задоволяването на някаква точно определена конкретна потребност. Такива стоки се различават само по **цени и натурални количества**, които са представени в **една и съща натурална мярка**. В математически смисъл еднородната съвкупност представлява **честотно разпределение** на елементите или единиците на едно крайно множество според значенията на една или повече характеристики (признаци). Когато не е изпълнено едното условие (единиците

* Професор, д.ик.н.; e-mail: emil_hristov_37@hotmail.com.

или елементите на множеството да са взаимозаменяеми) и/или не е изпълнено и второто условие (натуралното количество на единиците или елементите на множеството да са в една и съща натурална мярка), съвкупността е **разнородна**. По определение в икономиката тя е крайно множество от **разнородни стоки** или **подсъвкупности** на различни стоки, всяка от които е **еднородна съвкупност**. Определянето на една статистическа съвкупност обаче като еднородна или разнородна е най-напред **икономическа** и след това математическа задача. Причината е, че **видът на съвкупностите** се определя **предварително** и след това се съставя като еднородна или разнородна според някаква конкретна целесъобразност (Христов, 2016а, 2016б, 2017). Тъй като всяка отделна стока има индивидуална цена p и индивидуално количество q за една календарна година, за нея се предлага понятието „**елементарна съвкупност**“ в икономическата статистика. С това понятие всяка еднородна съвкупност на стоки може да се определи като крайно множество на **елементарни съвкупности от един и същ вид**, докато всяка разнородна съвкупност е крайно множество на **различни елементарни съвкупности (стоки)**. Възможно е една и съща стока (елементарна съвкупност) или група стоки от един и същ вид да участват по една целесъобразност в еднородна съвкупност, а по друга целесъобразност в разнородна съвкупност. Възможно е същата стока или група стоки да участват по различни целесъобразности в две различни еднородни или в две различни разнородни съвкупности. Посоченото различие между еднородните и разнородните съвкупности е фундаментално в икономическата статистика и е **гранична линия** между нея и приложената математика във всеки конкретен икономически анализ. Изводът от определенията на двата вида съвкупности, който може да се направи е, че само еднородната съвкупност на стоките има **икономически обоснована средна цена \bar{p}** от цените p_i на отделните стоки, както и **общо натурално количество на всички стоки Q** от натуралните количества q_i , които са в една и съща натурална мярка. Следователно обемът на продукцията P в паричен израз на всяка еднородна съвкупност стоки за една календарна година може да се изрази аналитично с двуфакторния мултипликативен модел $P = \bar{p} \times Q$, където стойностите на двата фактора \bar{p} и Q се получават с **агрегираните** данни за всички стоки от еднородната съвкупност, или $\bar{p} = \frac{\sum p_i \times q_i}{Q}$, където $Q = \sum q_i$ (Христов, 2016а). Когато цената p_i на всяка i -та отделна стока не е постоянна през цялата календарна година, точната стойност трябва да се определя като претеглена средна хронологична. С посочения модел се извършва адитивен факторен анализ на изменението в обема на продукцията $\Delta P = P_1 - P_0$ за всяка отчетна спрямо



друга (базисна) година. Разликата ΔP се обяснява с разликите на средните цени $\Delta \bar{p} = \bar{p}_1 - \bar{p}_0$ и на общите натурални количества на стоките $\Delta Q = Q_1 - Q_0$. Цялостната методика за този адитивен анализ на **продукцията** с агрегираните данни за \bar{p} и Q на стоките е изложена в моя предходна статия в списанието (Христов, 2016а). С нея се установяват **три ефекта** (два нетни - само от промяната на \bar{p} и от промяната на Q) и един съвместен ефект от еднопосочните едновременни увеличения или намаления на двата фактора, както и само **два нетни ефекта** от разнопосочните промени на тези фактори, **без съвместни ефекти**. Всички ефекти се получават с дискретната нечетна функция на математическия сигнум. Освен с **агрегираната информация** за средната цена \bar{p} и общото натурално количество Q за всички стоки от еднородна съвкупност, разликата в обемите на продукцията $\Delta P = P_1 - P_0$ може да се анализира и с **неагрегирана информация** за цените p_i и натуралните количества q_i на **отделните стоки**. Същата неагрегирана информация се отнася както за обема на продукцията от еднородни съвкупности, така и от разнородни съвкупности. За адитивния факторен анализ няма никакво значение **видът на съвкупността** (еднородна или разнородна), ако той се извършва по **отделни стоки**. Следователно методиките и съответните методи за адитивния факторен анализ на продукцията от съвкупности на стоки зависят не само от **вида на съвкупностите** - еднородна или разнородна, но и от **вида на информацията** - агрегирана или неагрегирана. Адитивният факторен анализ на изменението на продукцията $\Delta P = P_1 - P_0$ се извършва с агрегираната информация **само** с разликите на средните цени $\Delta \bar{p} = \bar{p}_1 - \bar{p}_0$ и на общите натурални количества на стоките $\Delta Q = Q_1 - Q_0$ **общо за всички стоки от сравняваните еднородни съвкупности**. Същата разлика на продукцията $\Delta P = P_1 - P_0$ може да се анализира и с **неагрегирана информация** с разликите $\Delta p_i = p_{i1} - p_{i0}$ и на натуралните количества $\Delta q_i = q_{i1} - q_{i0}$ за **отделните стоки на същите еднородни съвкупности** (Христов, 2016а). Двата адитивни факторни анализа на продукцията от еднородните съвкупности обаче са **различни**. Също за разлика от тези два анализа с еднородните съвкупности адитивният факторен анализ на продукцията от **разнородните съвкупности на стоки** може да се извършва **единствено и само с неагрегирана информация** с разликите на цените $\Delta p_i = p_{i1} - p_{i0}$ и на натуралните количества $\Delta q_i = q_{i1} - q_{i0}$ на **отделните различни стоки**, след като такива стоки не могат да имат икономически обоснована средна цена и общо натурално количество Q . Авторската методика за адитивния факторен анализ на продукцията **само от разнородните съвкупности на стоките** е публикувана в посочената статия за адитивния факторен анализ на продукцията от еднородните съвкупности на стоките (Христов, 2016а). От своя страна, методиката за

адитивния факторен анализ на продукцията на всяка отделна стока е публикувана в първата статия на автора за елементарния адитивен и индексен функционален анализ с дискретната нечетна функция на математическия сигнум (Христов, 2015). Освен посочените **две задължителни условия** за вида на статистическите съвкупности (еднородни или разнородни) и за вида на информацията (агрегирана или неагрегирана) има и още едно **препоръчително условие** само за неагрегираната информация на стоките. Тя може да бъде както с **групирани данни на стоките**, така и с **негрупирани данни за отделните стоки**. Във връзка с това е необходимо да се отбележи, че не е задължително разпределението или групирането на отделните стоки да бъде според величините на техните цени и натурални количества или по друг признак. Правилото е, че разпределението или групирането се определя от конкретната задача. В икономиката обаче, където се работи с парични средства, най-голямо значение за икономическия анализ има разпределението на стоките по цени и натурални количества. За съжаление, посочените съществени и важни условия за адитивния факторен анализ на продукцията от статистическите съвкупности на стоките липсват в икономическото образование по статистика.

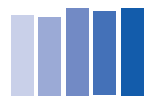
Предлаганата методика в настоящата статия е също за адитивен факторен анализ, но само на промяната или разликата на средните цени $\Delta \bar{p} = \bar{p}_1 - \bar{p}_0$. Този адитивен анализ може да бъде също с **неагрегирана информация** за отделните стоки, но само от **еднородни съвкупности на стоките**, защото само те имат **средни цени**. Следователно адитивният факторен анализ изпълнява и двете посочени условия. **Новата методика** за адитивен факторен анализ на средните цени представлява по-нататъшно развитие на предходната методика за адитивен анализ на продукцията с агрегираните данни за \bar{p} и Q на стоките и по тази причина в нея също се препоръчва групиране на стоките. За тази цел **натуралните количества на стоките q_i се превръщат в относителни дялове ω_i** спрямо тяхното общо натурално количество $Q = \sum q_i$ за всяка календарна година. Относителните дялове ω_i образуват **статистическата структура** на натуралните количества q_i , защото $\sum \omega_i = 1$. С относителните дялове $\omega_i = \frac{q_i}{Q}$ двуфакторният мултипликативен модел за продукцията $P = \bar{p} \times Q$ се превръща в трифакторен модел с групирани данни на стоките $P = \sum p_i \times \omega_i \times Q$. С този модел се изразява зависимостта на обема на продукцията P от трите фактора - цените p_i , относителните дялове ω_i на натуралните количества q_i и общото натурално количество на всички стоки Q . От него се извежда двуфакторният мултипликативен модел **само** за средната цена $\bar{p} = \sum p_i \times \omega_i$, от който следващият модел за адитивен факторен анализ на

промяната на тази средна цена е $\Delta\bar{p} = \bar{p}_1 - \bar{p}_0 = \frac{\sum p_{i1} \times q_{i1}}{Q_1} - \frac{\sum p_{i0} \times q_{i0}}{Q_0} = \sum p_{i1} \times \omega_{i1} - \sum p_{i0} \times \omega_{i0}$.

Решението на този модел се извършва също с дискретната нечетна функция на математическия сигнум. **Новото** в предлаганата методика е, че в общия случай с **четирите вида промени** (еднопосочни и разнопосочни) на двата фактора p_i и ω_i могат да се получат с дискретната нечетна функция на математическия сигнум всичко **осем различни решения** със сумарни ефекти от **всички стоки**. Всяко от тях е с **три сумарни ефекта** (два **нетни** - само от промените на p_i и промените на ω_i) и **един съвместен ефект** от еднопосочните едновременни увеличения или намаления на двата фактора. Компетентният читател ще забележи още тук, че осемте решения се различават методологично от решенията на адитивния факторен анализ на средните равнища в икономическото образование и в приложните изследвания на обществените науки в БАН. В точка 3 на настоящата статия са разкрити причините за техните различия. Получените осем решения са **подобни** на осемте решения с **общата методика** за адитивен факторен анализ на обема на продукцията от еднородните и разнородните съвкупности на стоките само с данните за цените p_i и натуралните количества q_i на **отделните стоки** (Христов, 2017). Тези осем решения с p_i и q_i се обясняват с **независимостта на продукциите** на отделните стоки $P_{oi} = p_i \times q_i$, откъдето тяхната сума е равна на обема на продукцията от цялата съвкупност на стоките, независимо дали тя е еднородна, или разнородна, $P_o = \sum P_{oi} = \sum p_i \times q_i$. По аналогичен начин и моделът за средната цена $\bar{p} = \sum p_i \times \omega_i$ с групирани данни на стоките се представя също като сума на **независимите величини** $p_i \times \omega_i$. Именно поради тази независимост адитивният факторен анализ на $\Delta\bar{p}$ се извършва за **всяка отделна стока** с дискретната нечетна функция на математическия сигнум. Получените ефекти от дадената факторна промяна представляват **приносите** на всяка i -та стока в тяхната сума на ефектите от всички стоки. По този начин за цялата еднородна съвкупност на стоките се съставят **три алгебрични суми** за трите различни ефекта. Отделно се сумират трите ефекта за всяка отделна стока. **Сумата** на трите отделни **сумарни ефекта** за всички стоки и **общата сума** на трите ефекта от всяка стока трябва да бъдат **равни** както помежду си, така и на разликата между двете средни цени $\Delta\bar{p} = \bar{p}_1 - \bar{p}_0$. Получените решения се очаква да съдържат в **общия случай** с четирите вида промени на двата фактора p_i и ω_i **минимални съвместни ефекти** от еднопосочните увеличения или намаления на факторите в сравнение с нетните ефекти

от **някои стоки**, защото останалите стоки в еднородната съвкупност с разнопосочните факторни промени са без каквито и да са съвместни ефекти.

Предлаганата методика за адитивен факторен анализ на средните цени с групирани данни за цените на стоките p_i и за относителните дялове ω_i на техните натурални количества q_i има своя отделна много важна цел, която е различна от задачите на другите методики за адитивен факторен анализ. Например за разлика от общата методика за влиянията на **промените на цените p_i и на натуралните количества q_i** на отделните стоки върху изменението на продукцията $\Delta P = P_1 - P_0 = \sum p_{i1}q_{i1} - \sum p_{i0}q_{i0}$, която е публикувана в моята предходна статия, тук адитивният факторен анализ е за влиянията също на **промените на цените p_i** , но не на промените на натуралните количества q_i , а на **техните относителни дялове ω_i** върху разликата на двете средни цени $\Delta \bar{p} = \bar{p}_1 - \bar{p}_0 = \sum p_{i1} \times \omega_{i1} - \sum p_{i0} \times \omega_{i0}$ (Христов, 2017). Това са **два различни анализа**, защото относителните дялове $\omega_i = \frac{q_i}{Q}$ зависят не само от натуралните количества q_i , но и от тяхната сума $\sum q_i = Q$. От тази зависимост разликите $\Delta \omega_i = \omega_{i1} - \omega_{i0}$ за някои стоки **не съответстват** по алгебричен знак и величина на разликите $\Delta q_i = q_{i1} - q_{i0}$. От адитивния факторен анализ на разликата $\Delta \bar{p}$ се получават **три** много важни сумарни ефекта от ефектите за отделните стоки (прирасти и/или намаления на базисната цена \bar{p}_0). Първият сумарен ефект е **ценовият** само от промените на цените p_i на отделните стоки, вторият сумарен ефект е **структурният** само от структурните промени на теглата q_i на стоките чрез разликите на техните относителни дялове ω_i , докато третият сумарен ефект е **съвместният** само от еднопосочните промени (едновременни увеличения или намаления) на двата фактора p_i и ω_i на отделните стоки. Поради това методиката за адитивния факторен анализ на разликата на средните цени $\Delta \bar{p}$ е **универсална** за решенията на всички задачи с този анализ на всякакви средни равнища. От трите сумарни ефекта специално значение има **общият структурен ефект** (в т.ч. нетният и неговата пропорционална част от съвместния ефект), който не може да се измери коректно по никакъв друг начин. Следователно методиката може да се прилага не само в икономиката, но и във всички други области на живота, където се извършват анализи с еднородни съвкупности. Във връзка с посочената универсалност на предлаганата методика за адитивния факторен анализ на средни равнища тя е можела да бъде изведена отдавна в демографската статистика не само с индуктивната обобщаваща логика на автора или със знаковата функция на математическия сигнум, но и с **теорията на вероятностите**. Демографската статистика е била включена след Втората световна война и като



социална и икономическа за трудовите ресурси (работната сила) и за заетите в икономическото образование и обществените науки у нас. В аналитично отношение обаче тя е **пряко свързана** с теорията на вероятностите, от която могат да се изведат еднозначните решения на адитивния факторен анализ на всякакви средни равнища от еднородните статистически съвкупности. Защо това не е направено досега, е абсолютно необяснимо. Именно за нуждите на всички **неикономически статистики** е изложен в следващата точка 1 на настоящата статия адитивният факторен анализ с теорията на вероятностите. С него се анализира разликата на средни равнища (общите коефициенти за смъртност на населението за базисна и отчетна година) от промените на неговата смъртност по възраст и на относителните дялове на неговия средногодишен брой по възраст за двете сравнявани години.

1. Адитивен факторен анализ на средните цени на стоки от еднородни съвкупности и обобщение за средните равнища (показатели за интензивност) от всякакви еднородни статистически съвкупности

Една от най-разпространените практики на адитивния факторен анализ е на промяната на средната цена от промените на **групирани данни** на стоките според техните цени и натурални количества (честотни разпределения) за две сравнявани години (базисна и отчетна). Адитивен анализ може да се извърши и със същите, но **негрупирани данни** на стоките, защото с тях може също да се пресмята същата средна цена. Групираните данни обаче са много полезни и указателни за икономическия анализ и от анализатора зависи видът на групирането. Една от най-полезните групировки е по величина на цените на отделните стоки, но тя може да бъде и с други групи стоки по друг признак, който също зависи от решението на анализатора за броя на групите и за различния брой на стоките в отделните групи.

Изходната формула за средната цена \bar{p} от всички данни на стоките за техните цени p_i и натурални количества q_i (теглата) е $\bar{p} = \frac{\sum p_i \times q_i}{\sum q_i} = \frac{\sum p_i \times q_i}{Q}$, където $Q = \sum q_i$ е общото натурално количество на всички стоки в една и съща натурална мярка, i е индексът (номерът) на всяка отделна стока 1, 2, ..., n (Гатев, 1995).

Ако всички **n** стоки са групирани в по-малък брой **m** групи със своите групови цени \bar{p}_j и натурални количества q_j в отделните групи (еднородни съвкупности като подсъвкупности на разнородната съвкупност), могат за удобство на анализа да се направят заместванията $\bar{p}_j = p_i$ и $q_j = q_i$. Никога обаче не трябва да се забравя, че съответните i -та са за отделните **еднородни подсъвкупности на стоките** (Христов, 2017).

Решението на адитивния факторен анализ е на разликата на двете средни претеглени цени за базисната и отчетната година:

$$\Delta \bar{p} = \bar{p}_1 - \bar{p}_0 = \frac{\sum p_{i1} \times q_{i1}}{Q_1} - \frac{\sum p_{i0} \times q_{i0}}{Q_0}.$$

С цел да се елиминира влиянието на промяната на **третия фактор** - общото натурално количество на стоките Q , вторият фактор - натуралните количества на отделните стоки q_i , се превръща в относителни дялове $\omega_i = \frac{q_i}{Q}$, с които се претеглят цените, или първият фактор. По този начин вторият фактор q_i се превръща в **структурен фактор** ω_i , откъдето двуфакторният модел на разликата на средните цени получава крайния израз с ω_i :

$$\Delta \bar{p} = \frac{\sum p_{i1} \times q_{i1}}{\sum q_{i1}} - \frac{\sum p_{i0} \times q_{i0}}{\sum q_{i0}} = \frac{\sum p_{i1} \times q_{i1}}{Q_1} - \frac{\sum p_{i0} \times q_{i0}}{Q_0} = \sum p_{i1} \omega_{i1} - \sum p_{i0} \omega_{i0}.$$

В общия случай всяка стока се характеризира с един от следните **четири вида** факторни промени на цената p_i и относителния дял ω_i на натуралното количество q_i :

- едновременни увеличения на двата фактора $\Delta p_i > 0$ и $\Delta \omega_i > 0$;
- едновременни намаления на двата фактора $\Delta p_i < 0$ и $\Delta \omega_i < 0$;
- разнопосочни факторни промени (единият фактор се е увеличил, а другият е намалял) $\Delta p_i > 0$ и $\Delta \omega_i < 0$ или $\Delta p_i < 0$ и $\Delta \omega_i > 0$.

От тези четири възможни промени на двата фактора произлизат следните **ефекти от всяка i -та стока** с дискретната нечетна функция на математическия сигнум:

$\Delta p, \omega_i = \Delta p_i \times \omega_{imin}$ - **нетен ценови ефект** (прираст или намаление на средната базисна цена \bar{p}_0) **само** от промяната (увеличението или намалението) на цената $\Delta p_i = p_{i1} - p_{i0}$, където ω_{imin} е **по-малкият относителен дял** ω_i на i -тата стока от базисната или от отчетната година;

$\Delta \omega, p_i = \Delta \omega_i \times p_{imin}$ - **нетен структурен ефект** (прираст или намаление на средната базисна цена \bar{p}_0) **само** от промяната (увеличението или намалението) на относителния дял $\Delta \omega_i = \omega_{i1} - \omega_{i0}$, където p_{imin} е **по-малката цена** на i -тата стока от базисната или от отчетната година;

$h_i \Delta p_i \Delta \omega_i$ - **съвместен ефект** (прираст или намаление на средната базисна цена \bar{p}_0) **само** от еднопосочни съвместни промени на цената Δp_i и на относителния дял $\Delta \omega_i$ на i -тата стока, където $h_i = +1, 0$ или -1 са трите дискретни стойности на нечетната функция на математическия сигнум. С тях се определя **наличието** на съвместен ефект със съответния алгебричен знак или **отсъствието** на такъв ефект (Христов, 2015).

При едновременни увеличения на двата фактора $\Delta p_i > 0$ и $\Delta \omega_i > 0$, $h_i = +1$, откъдето съвместният ефект е положителна величина $\Delta p_i \Delta \omega_i > 0$. При едновременни намаления на факторите $\Delta p_i < 0$ и $\Delta \omega_i < 0$, $h_i = -1$, същият ефект е отрицателна величина $\Delta p_i \Delta \omega_i < 0$, докато при разнопосочните факторни промени $\Delta p_i > 0$ и $\Delta \omega_i < 0$ или $\Delta p_i < 0$ и $\Delta \omega_i > 0$, $h_i = 0$, **няма** съвместни ефекти.

Обобщено, само от еднопосочните факторни промени има **три ефекта** (два нетни) и един съвместен с еднакви положителни или отрицателни алгебрични знаци, докато от разнопосочните факторни промени (единият фактор се е увеличил, а другият е намалял) има само **два нетни ефекта** с различни алгебрични знаци, **без съвместен ефект**.

Подобно на методиките за адитивен факторен анализ на обемните резултативни величини (продукцията) с дискретната функция на математическия сигнум и в настоящата методика за адитивен анализ на средните равнища (цени) факторните промени на всяка стока се отчитат според **концепцията за тяхната взаимозависимост** (The Oxford Paperback Dictionary, 1994). С тази концепция промяната на всеки фактор се **съобразява** с едновременната промяна на другия фактор. Всъщност това е „тайната“ на всеки верен адитивен факторен анализ, според която промените на всички фактори трябва да се отчитат **едновременно**, а не поотделно. Това строго логическо и математическо условие се изпълнява именно от дискретната нечетна функция на математическия сигнум. С нея всяка промяна на единия фактор трябва да се умножава с **по-малката стойност** на другия фактор. Ако всички промени на единия фактор се умножават със стойностите на другия фактор **само** от базисната година или **само** със стойностите от отчетната година, така както се предлага за адитивния факторен анализ във всички учебници и ръководства по теория на статистиката и за всички отраслови статистики, се получават **винаги неверни ефекти за една част на стоките**, а оттам и за цялата тяхна съвкупност! Тези неверни ефекти произлизат както от еднопосочните факторни промени, така и от разнопосочните. Те са показани като неверни в точка 3 на статията.

На следващ етап се сумират **отделно ефектите от всеки вид** (ценовите, структурните и съвместните) **за всички стоки** на еднородната съвкупност:

$$\begin{aligned}\sum \Delta p_i \omega_{imin} &= \sum \Delta p, \omega_i = \Delta \bar{p}, p, \\ \sum \Delta \omega_i p_{imin} &= \sum \Delta \omega, p_i = \Delta \bar{p}, \omega, \\ \sum h_i \Delta p_i \Delta \omega_i &= \Delta \bar{p}, p \omega.\end{aligned}$$

Алгебричната сума на трите сумарни ефекта със своите алгебрични знаци е равна на разликата (прирастът или намалението) на средната цена:

$$\Delta \bar{p}, p + \Delta \bar{p}, \omega + \Delta \bar{p}, p \omega = \Delta \bar{p}.$$

Или обобщено, от адитивния факторен анализ на средните цени се получават в общия случай от четирите вида факторни промени **три отделни (независими) сумарни ефекта** с дискретната нечетна функция на математическия сигнум:

– $\Delta \bar{p}, p$ е **общият (сумарен) ценови ефект** (салдо) само от **преобладаваща сума** на положителни или отрицателни ценови ефекти $\Delta p, \omega_i$ от промените на цените на отделните стоки $\Delta p_i = p_{i1} - p_{i0}$;

– $\Delta \bar{p}, \omega$ е **общият (сумарен) структурен ефект** (салдо) само от **преобладаваща сума** на положителни или отрицателни структурни ефекти $\Delta \omega, p_i$ от промените на относителните дялове $\Delta \omega_i = \omega_{i1} - \omega_{i0}$ на натуралните количества на стоките; и

– $\Delta \bar{p}, p \omega$ е **общият (сумарен) съвместен ефект** (салдо) само от **преобладаваща сума** на положителни или отрицателни съвместни ефекти $h_i \Delta p_i \Delta \omega_i$ от **еднопосочни съвместни промени** на цените Δp_i и на относителните дялове $\Delta \omega_i$ на натуралните количества на стоките. Необходимо е винаги да се има предвид, че общият (сумарен) съвместен ефект се получава **само** от онези стоки, които имат съвместни ефекти, защото не всички стоки може да са с еднопосочни факторни промени на цените Δp_i и на относителните дялове $\Delta \omega_i$.

Алгебричният знак на всеки от трите общи (сумарни) ефекта зависи от **величините на** положителните и отрицателните ефекти на отделните стоки. Същите ефекти представляват съответните **приноси** на всяка стока във всеки сумарен ефект, както и в разликата на двете средни цени $\Delta \bar{p} = \bar{p}_1 - \bar{p}_0$. Тази разлика може да се

представи и като **сума на общите приноси** $p_{i1}\omega_{i1} - p_{i0}\omega_{i0}$ на всички стоки, всеки принос от които е равен на **сумата** на трите ефекта на i -тата стока: $\Delta p, \omega_i + \Delta \omega, p_i + h_i \Delta p_i \Delta \omega_i = p_{i1}\omega_{i1} - p_{i0}\omega_{i0}$, откъдето за всички стоки прирастът или намалението на средната цена е $\sum(p_{i1}\omega_{i1} - p_{i0}\omega_{i0}) = \Delta \bar{p}$.

Новото в изложената методика за тези решения е, че за разлика от известните общо четири различни решения на авторите от икономическото образование те са **осем**, всяко от които е с трите сумарни ефекта. В шест от тези решения два от трите сумарни ефекта (нетни или нетен и съвместен) са с **еднакви** положителни или отрицателни алгебрични знаци, докато третият ефект (нетен или съвместен) е с **обратен** отрицателен или положителен знак. В останалите две решения и трите ефекта са с еднакви положителни или отрицателни алгебрични знаци. Или аналитично, **осемте еднозначни** (единствени) верни и точни решения са следните:

1. $\Delta \bar{p}, p > 0$, $\Delta \bar{p}, \omega < 0$ и $\Delta \bar{p}, p\omega > 0$
2. $\Delta \bar{p}, p < 0$, $\Delta \bar{p}, \omega > 0$ и $\Delta \bar{p}, p\omega < 0$
3. $\Delta \bar{p}, p > 0$, $\Delta \bar{p}, \omega < 0$ и $\Delta \bar{p}, p\omega < 0$
4. $\Delta \bar{p}, p < 0$, $\Delta \bar{p}, \omega > 0$ и $\Delta \bar{p}, p\omega > 0$
5. $\Delta \bar{p}, p > 0$, $\Delta \bar{p}, \omega > 0$ и $\Delta \bar{p}, p\omega < 0$
6. $\Delta \bar{p}, p < 0$, $\Delta \bar{p}, \omega < 0$ и $\Delta \bar{p}, p\omega > 0$
7. $\Delta \bar{p}, p > 0$, $\Delta \bar{p}, \omega > 0$ и $\Delta \bar{p}, p\omega > 0$
8. $\Delta \bar{p}, p < 0$, $\Delta \bar{p}, \omega < 0$ и $\Delta \bar{p}, p\omega < 0$.

Частният случай само с двата нетни факторни ефекта $\Delta \bar{p}, p$ и $\Delta \bar{p}, \omega$ е извънредно рядък.

Ако внимателният читател е запознат с **общата методика** на автора за **изменението на продукцията** от еднородните и разнородните съвкупности на стоките с данните за цените на отделните стоки p_i и на техните натурални количества q_i , ще установи, че и адитивният анализ на **средните цени** се извършва със **същата методика** (Христов, 2017). Адитивните анализи на изменението на продукцията $\Delta P = P_1 - P_0 = \sum p_{i1} \times q_{i1} - \sum p_{i0} \times q_{i0}$ и на разликата на средните цени $\Delta \bar{p} = \bar{p}_1 - \bar{p}_0 = \sum p_{i1} \times \omega_{i1} - \sum p_{i0} \times \omega_{i0}$ показват, че първият фактор (цените p_i) е един и същ в двата анализа. Вторият фактор обаче, q_i за първия анализ и ω_i за втория анализ, е **различен**, защото $\omega_i = \frac{q_i}{Q}$ зависят не само от q_i , но и от Q . Оттук и разликите $\Delta \omega_i = \omega_{i1} - \omega_{i0}$ зависят не само от разликите $\Delta q_i = q_{i1} - q_{i0}$, но и от отношението $\frac{Q_1}{Q_0}$. Следователно двата

адитивни факторни анализа на разликите ΔP и $\Delta \bar{p}$ са **различни** поради различието на вторите фактори, от които се получават различни ефекти - едните са увеличения и/или намаления на обемна резултативна величина (продукцията), а другите са увеличения и/или намаления на средно равнище (средната цена). Или независимо от еднаквия логически и аналитичен подход за адитивния факторен анализ двата анализа са **различни**.

В общия случай само една част от стоките са с еднопосочни промени (увеличения и намаления) на двата фактора, от които произлизат съвместните ефекти. В сравнение с нетните ефекти те са много **по-малки величини**, защото са алгебрични суми на положителни съвместни ефекти за едни стоки и отрицателни съвместни ефекти за други стоки. Останалата част от стоките са с разнопосочни факторни промени, от които няма съвместни ефекти.

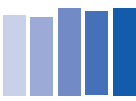
Освен с индуктивната обобщаваща логика и нейния аналитичен аналог - дискретната функция на математическия сигнум, еднозначните решения на адитивния факторен анализ на средните равнища могат да се изведат и с приложението на **теорията на вероятностите** в демографската статистика. Според мен, след като тази теория е важна част от теоретичната математика, не би следвало приложенията на дискретните статистически методи в демографията, трудовата и социалната, както и в застрахователната статистика, да имат каквито и да било теоретико-методологични проблеми. Именно във връзка с демографията най-точните статистически показатели за интензивност на демографските явления и процеси са **вероятности** за възникването на демографските събития, като раждане, умирање, заселване, изселване и други, а населението по пол и възраст са **средите** (най-точните екстензивни показатели), от които произлизат тези събития. В този смисъл вероятностите и средите са **типичен пример** за дискретни (прекъснати) променливи, защото се съставят за всяка отделна година. На тази основа може да се изведе **точният адитивен факторен анализ**, например на разликите в броя на умрелите мъже и жени (разлики на обемни резултативни величини) по възраст $\Delta D_x = D_{xm} - D_{xf}$ за даден календарен период (година), където x е точната възраст в години, D_{xm} и D_{xf} са броят на умрелите мъже и броят на умрелите жени. Същите разлики между мъжете и жените зависят от разликите на вероятностите за умирање по възраст на мъжете и жените $\Delta q_x = q_{xm} - q_{xf}$ и разликите в броя на мъжете и жените на точните възрасти x години $\Delta S_x = S_{xm} - S_{xf}$ или средите, от които произлизат умрелите. По мое мнение този анализ не може да бъде проблем за математиците, които преподават теория на вероятностите и демографска статистика на икономистите статистици. Всички статистици и

математици, които се занимават по някакъв повод с демография, знаят, че в началните детски възрасти вероятностите за умирање на момчетата са **по-големи** от вероятностите за умирање на момичетата и че броят на момчетата е също **по-голям** от броя на момичетата поради техния по-голям брой при раждане. Ако се извърши адитивен факторен анализ с теорията на вероятностите на разликата на по-големия брой на умрелите момчета и по-малкия брой на умрелите момичета на дадена детска възраст x години, се получават **три ефекта** - **два нетни** (прирасти на умрели момчета) и **един съвместен** също прираст на умрели момчета. Сумата на трите ефекта (прирасти) е точно равна на разликата на умрелите момчета и момичета ΔD_x . Същата, но **обратна разлика** обаче може да се анализира също с теорията на вероятностите, защото за тази теория е без значение алгебричният знак на разликата. Тогава математиците и статистиците сигурно ще установят, че обратната (отрицателна) разлика - намаление на умрелите момичета, се състои от същите по абсолютна стойност два нетни ефекта и съвместния ефект както от анализа на първата (положителна) разлика за умрелите момчета. Трите ефекта за умрелите момичета обаче ще са с обратни алгебрични знаци, тъй като са две нетни намаления и едно съвместно намаление на умрели момичета. Трите ефекта с еднаквите положителни алгебрични знаци за момчетата и отрицателни за момичетата се получават от **еднопосочните различия** на двата фактора $\Delta q_x = (q_{xm} - q_{xf}) > 0$ и $\Delta S_x = (S_{xm} - S_{xf}) > 0$ за момчетата или $\Delta q_x = (q_{xm} - q_{xf}) < 0$ и $\Delta S_x = (S_{xm} - S_{xf}) < 0$ за момичетата. Верните нетни ефекти **само** от разликата на единия фактор и **само** от разликата на другия фактор се получават чрез теорията на вероятностите, когато разликата на всеки фактор се умножи с **по-малката стойност** на другия фактор за момчетата или за момичетата както с дискретната функция на математическия сигнум. Ако разликата на всеки фактор се умножи с по-голямата стойност на другия фактор за момчетата или за момичетата, се получават **неверни ефекти** за някои възрасти, с които теорията на вероятностите няма нищо общо. Такива по-големи стойности на другия фактор участват в анализа, когато предварително е решено да се използват неговите стойности **само** от базисната или **само** от отчетната година.

Освен трите ефекта с еднаквите алгебрични знаци за началните детски възрасти ефектите от факторните различия между мъжете и жените за следващите по-високи възрасти са с различни знаци поради по-голямата смъртност на мъжете $\Delta q_x = (q_{xm} - q_{xf}) > 0$ и по-малкия техен брой $\Delta S_x = (S_{xm} - S_{xf}) < 0$. По изключение за някои възрасти може да се срещне и обратният случай на разнопосочните промени с $\Delta q_x =$

$(q_{xm} - q_{xf}) < 0$ и $\Delta S_x = (S_{xm} - S_{xf}) > 0$. Според посоченото правило за измерване на верните ефекти от всички разнопосочни факторни различия произлизат само **два нетни ефекта** с различни алгебрични знаци, **без съвместен ефект**. Ако разликата на всеки фактор се умножи с по-голямата стойност на другия фактор за мъжете или за жените, се получават **два брутни неверни ефекта**. Единият от тях ще съдържа положителния верен нетен ефект и един също положителен, но **фиктивен** (несъществуващ) съвместен ефект. **Едновременно** обаче другият брутен ефект ще съдържа отрицателния верен нетен ефект и същия по размер **фиктивен**, но отрицателен съвместен ефект! Естествено, че теорията на вероятностите няма никога да допусне такива безсмислици. Причината е известна. Двата едновременни фиктивни ефекта през отчетната година означават логически и демографски, че през тази година е **имало допълнително увеличение на умрелите мъже** в размер на положителния фиктивен съвместен ефект и едновременно е **имало допълнително намаление на умрелите мъже** в същия размер от отрицателния фиктивен съвместен ефект! Тъй като двата съвместни ефекта са **едни и същи умрели хора (мъже)**, всеки може да запита действително ли са умрели **през базисната година**, за да бъдат **допълнително намаление** на умрелите мъже през отчетната година и как така са умрели **също и през отчетната година**, за да бъдат **допълнително увеличение** на умрелите? Според авторите на такива сбъркани демографски анализи излиза, че **едни и същи хора** са умрели **едновременно** през базисната и през отчетната година! Възможно е да има обаче и още едно сбъркано обяснение на двата фиктивни съвместни ефекта. Ако **същите хора (мъже)** са умрели само през отчетната година като **допълнително увеличение на умрелите мъже**, те трябва **едновременно и да са неумрели** през същата отчетна година, за да **анулират** допълнителното увеличение. Това означава, че веднага след своята смърт през отчетната година те **моментално са оживели!** Истината обаче е само една. Според нея при разнопосочните факторни промени **няма никакви фиктивни ефекти от адитивния факторен анализ**. Тази истина може да се обоснове логически както с предходния текст, така и както се извежда аналитично с дискретната функция на математическия сигнум и теорията на вероятностите!

Със същия метод за адитивен факторен анализ на групирани данни по възраст може да се анализира и разликата на два **общи** коефициента на смъртността на населението за два календарни периода (базисна и отчетна година). Двата коефициента, с които работи текущата практика, са известни като **общи**, защото се отнасят за цялото население от всички възрасти. Те представляват две претеглени средни равнища на



смъртността за базисната година $\bar{m}_0 = \sum m_{i0}\omega_{i0}$ и за отчетната година $\bar{m}_1 = \sum m_{i1}\omega_{i1}$. В тях m_{i0} и m_{i1} са коефициентите за смъртност на населението на всяка отделна възраст, а ω_{i0} и ω_{i1} са относителните дялове на средногодишното население на същата възраст от двете негови **възрастови структури** за базисната и отчетната година.

Според теорията на вероятностите всеки ефект трябва да се определя с произведението на промяната на единия фактор с **по-малката стойност** на другия фактор от базисната или от отчетната година. Според това строго правило е **логически недопустимо** промените на коефициентите на смъртността на отделните възрасти $\Delta m_i = m_{i1} - m_{i0}$ да се умножават с относителните дялове **само** от базисната година ω_{i0} **или** с относителните дялове **само** от отчетната година ω_{i1} , както предварително се решава в практиката по усмотрение на анализаторите. По същия начин е **логически недопустимо** промените на относителните дялове $\Delta\omega_i = \omega_{i1} - \omega_{i0}$ на населението на отделните възрасти да се умножават с коефициентите на смъртността **само** от базисната година m_{i0} **или** с коефициентите на смъртността **само** от отчетната година m_{i1} , както също предварително се решава от анализаторите. Логическата недопустимост произлиза от факта, че за някои детски и млади възрасти относителните дялове от двете сравнявани години са $\omega_{i0} > \omega_{i1}$, а за повечето следващи по-големи средни и старчески възрасти относителните дялове са $\omega_{i0} < \omega_{i1}$ поради остаряване на населението. Дори само от формална гледна точка **не е възможно** относителните дялове на едната структура ω_{i0} да бъдат за **всички възрасти по-големи или по-малки** от относителните дялове ω_{i1} на другата структура. Не е възможно да се допускат тези неравенства за всички възрасти, защото се нарушава строгото математическо условие, че сумата на положителните структурни промени $\sum \Delta\omega_i > 0$ трябва да бъде равна по абсолютна стойност на сумата на отрицателните структурни промени $\sum |\Delta\omega_i < 0|$ или $\sum |\Delta\omega_i > 0| = \sum |\Delta\omega_i < 0|$. Подобни фактически неравенства се наблюдават и при коефициентите на повъзрастовата смъртност. Повечето детски и млади възрасти са с неравенства $m_{i0} > m_{i1}$, а повечето средни и старчески възрасти са с обратни неравенства $m_{i0} < m_{i1}$. За тези промени на повъзрастовата смъртност обаче не важи математическото правило $\sum |\Delta m_i > 0| = \sum |\Delta m_i < 0|$, защото сумата на коефициентите m_i не е равна на 1 както сумата на относителните дялове $\sum \omega_i = 1$. Изводът от фактическите неравенства на двата фактора m_i и ω_i е, че за **някои възрасти** (млади и стари) **има еднопосочни** факторни промени $\omega_{i0} > \omega_{i1}$ и $m_{i0} > m_{i1}$, от които произлизат отрицателни съвместни ефекти $\Delta m_i \Delta \omega_i < 0$. За **повечето възрасти**, но не за всички, също има **еднопосочни** факторни промени $\omega_{i0} < \omega_{i1}$ и $m_{i0} < m_{i1}$, от които

произлизат положителни съвместни ефекти $\Delta m_i \Delta \omega_i > 0$. За последната (трета) група възрасти факторните промени са **разнопосочни** $\omega_{i0} > \omega_{i1}$ и $m_{i0} < m_{i1}$, както и $\omega_{i0} < \omega_{i1}$ и $m_{i0} > m_{i1}$. От такива факторни промени **няма** съвместни ефекти. Следователно е **недопустимо** промените на всеки фактор да се умножават за всички възрасти със стойностите на другия фактор **само** от базисната или **само** от отчетната година. Точно това се прави обаче с методите за адитивен факторен анализ в икономическото образование и в обществените науки у нас (икономика, демография и социология). Всъщност с тези методи могат да се решават **условни задачи**, например какви ефекти биха се получили от промените на единия фактор, ако се бяха запазили и през отчетната година същите базисни стойности на другия фактор. По мое мнение едва ли някой би оспорил условното познавателно значение на такива решения. С тях обаче двуфакторният модел на анализа се превръща в **два условни еднофакторни модела**, с които никога не може да се намери еднозначното, единствено вярно логично и математически точно решение от **едновременните промени** на двата фактора. Например при разнопосочните факторни промени за някои възрасти с $\Delta m_i > 0$ и $\Delta \omega_i < 0$ или $\Delta m_i < 0$ и $\Delta \omega_i > 0$ ефектите ще **бъдат винаги неверни**, ако промените на всеки фактор се умножават **неправилно** с по-големите стойности на другия фактор. В такива случаи за всяка такава i -та възраст се получават **два брутни неверни ефекта** $\Delta m, \omega_i$ и $\Delta \omega, m_i$. Всеки от тях ще съдържа **верен нетен ефект** и допълнителен **фиктивен (несъществуващ) съвместен ефект** със същия алгебричен знак на нетния ефект. Двата фиктивни ефекта са равни по абсолютна стойност, но са с различни алгебрични знаци. Сумата на двата брутни ефекта е точно равна на разликата на смъртността за същата i -та възраст $\Delta m_i = m_{i1} - m_{i0}$, защото двата фиктивни ефекта с разнопосочните алгебрични знаци взаимно се анулират. Брутните ефекти обаче са **неверни**, защото включват освен двата верни нетни ефекта още и двата безсмислени и неверни фиктивни ефекта. Според тях през една и съща отчетна година е **имало допълнително увеличение на смъртността** от положителния фиктивен съвместен ефект и **едновременно е имало допълнително намаление на смъртността** от отрицателния фиктивен съвместен ефект, което било равно по размер на допълнителното увеличение! Човек може много да се чуди на такава логика, но сериозните последици от нея са, че тя произвежда **неверни брутни ефекти**, защото факторните промени за всяка i -та възраст се отчитат **поотделно** една след друга, а **не едновременно**. По този начин и сумата на ефектите от всички възрасти е точно равна на разликата на двата общи коефициента на смъртността $\Delta \bar{m} = \bar{m}_1 - \bar{m}_0$, независимо че **ефектите за някои възрасти ще бъдат винаги неверни!** Следователно с теорията на

вероятностите се получават **същите верни ефекти и решения** на адитивния факторен анализ както с дискретната нечетна функция на математическия сигнум.

Представените решения за средните цени на стоките са проверени с примери. Редът на примерите е според тяхната последователност в предходни мои публикации (Христов, 2016а, 2016б). Всеки два последователни примера са с едни и същи данни, но с разменени места за базисната и отчетната година, откъдето техните решения са **взаимнообратими**. Те съдържат едни и същи ефекти по абсолютна стойност, но с **обратни** алгебрични знаци. Примерите в предходните публикации, които са с данни за цените p_i и натуралните количества на стоките q_i , са поместени в Приложение 1 - 6, докато средните цени в настоящата статия с цените p_i и относителните дялове ω_i на натуралните количества на стоките са представени в таблици 1 - 6. В следващата точка 2 са решени примерите, за които табличните данни са с относителните дялове ω_i .

2. Приложения на метода за адитивен факторен анализ на средните цени от групирани данни на стоките

Примерите за отделните осем случая с трите ефекта $\Delta\bar{p}, p$, $\Delta\bar{p}, \omega$ и $\Delta\bar{p}, p\omega$ от предходната точка 1 са съставени с четири и шест стоки. Всяка от тях се характеризира със средни хронологични цени \bar{p}_{i0} и \bar{p}_{i1} , които за удобство са заменени със символите p_{i0} и p_{i1} за базисна и отчетна година. Освен с цените отделните стоки се характеризират и с относителните си дялове ω_{i0} и ω_{i1} на техните натурални количества q_{i0} и q_{i1} . Същите стоки могат да се интерпретират и като четири или шест групи стоки (подсъвкупности) на еднородните съвкупности на стоките.

2.1. Адитивен факторен анализ на средните цени от групирани данни на стоки с ефектите $\Delta\bar{p}, p > 0$, $\Delta\bar{p}, \omega < 0$ и $\Delta\bar{p}, p\omega > 0$

Първият случай е с положителен ценови ефект $\Delta\bar{p}, p > 0$, отрицателен структурен ефект $\Delta\bar{p}, \omega < 0$ и положителен съвместен ефект $\Delta\bar{p}, p\omega > 0$. Входните данни за примера на този случай са представени в Приложение 1 и в табл. 1.

1. Адитивен факторен анализ на средните цени на стоките

Стоки	Относителни дялове		Ефекти от промени на цените			Ефекти от промени на относителните дялове			Съвместни ефекти	$p_{i1}\omega_{i1} - p_{i0}\omega_{i0}$
	ω_{i0}	ω_{i1}	Δp_i	ω_{im}	$\Delta p_i \omega_{im}$	$\Delta \omega_i$	p_{im}	$\Delta \omega_i p_{im}$		
i	1	2	3	4	5 = 3x4	6 = 2-1	7	8 = 6x7	9	10 = 5+8+9
А	0.2000	0.2777	+40	0.2000	8.0000	0.0777	40	3.1111	3.1111	14.2222
Б	0.2500	0.1667	-10	0.1667	-1.6667	-0.0833	50	-4.1667	-0.8333	-6.6667
В	0.2500	0.1667	+40	0.1667	6.6667	-0.0833	50	-4.1667	-	2.5000
Г	0,3000	0.3889	-10	0.3000	-3.0000	0.0889	50	4.4444	-	1.4444
Общо	1.0000	1.0000	X	X	+10	0	X	-0.7778	2.2778	11.5

От данните в табл. 1 се установява, че средната цена на стоките за базисната година е $\bar{p}_0 = \frac{\sum p_{i0}q_{i0}}{Q_0} = \sum p_{i0}\omega_{i0} = 53.5$ хил. лв., а за отчетната година е $\bar{p}_1 = \frac{\sum p_{i1}q_{i1}}{Q_1} = \sum p_{i1}\omega_{i1} = 65.0$ хил. лв., откъдето средната базисна цена \bar{p}_0 се е увеличила с $\Delta\bar{p} = \bar{p}_1 - \bar{p}_0 = 65.0 - 53.5 = 11.5$ хил. лева (табл. 1, колона 10). Това увеличение се състои от следните ефекти:

– ценовият $\Delta\bar{p}, p$ само от промените на цените на стоките е увеличение на средната базисна цена с $\Delta\bar{p}, p = \sum \Delta p_i \omega_{imin} = 10$ хил. лева (табл. 1, колона 5). Той се образува от общия по-голям положителен ефект 14.6667 хил. лв. от преобладаващото влияние на увеличените цени на първата и третата стока с общия по-малък отрицателен ефект – 4.6667 хил. лв. от по-слабите намаления на цените на втората и четвъртата стока. Приносите на отделните стоки за ефекта $\Delta\bar{p}, p$ са отделните ценови ефекти за всяка стока в колона 5 на табл. 1.

– структурният ефект $\Delta\bar{p}, \omega$ само от промените на относителните дялове на натуралните количества на стоките е минималното намаление на средната базисна цена с $\Delta\bar{p}, \omega = \sum \Delta \omega_i p_{imin} = -0.7778$ хил. лева (табл. 1, колона 8). Това минимално намаление произлиза от общия малко по-голям отрицателен ефект –8.3333 хил. лв. от преобладаващото влияние на намаленията на относителните дялове на втората и третата стока, и от общия по-малък положителен ефект 7.5555 хил. лв. от увеличенията на относителните дялове на първата и четвъртата стока. Приносите на отделните стоки за ефекта $\Delta\bar{p}, \omega$ са отделните структурни ефекти за всяка стока в колона 8 на табл. 1.

– съвместният ефект $\Delta\bar{p}, p\omega$ само от еднопосочните съвместни промени на цените на стоките и на относителните дялове на техните натурални количества е увеличението на средната базисна цена с $\Delta\bar{p}, p\omega = \sum h_i \Delta p_i \Delta \omega_i = 2.2778$ хил. лева (табл. 1, колона 9). Това увеличение се получава от по-големия положителен ефект 3.1111 хил. лв. на първата стока и от по-малкия отрицателен съвместен ефект -0.8333 хил. лв. на втората стока. Другите две стоки (третата и четвъртата) са само с разнопосочни факторни промени, от които няма съвместни ефекти.

Алгебричната сума на измерените три ефекта е точно равна на увеличението на средната базисна цена с $\Delta\bar{p} = 11.5$ хил. лв., или $\Delta\bar{p}, p + \Delta\bar{p}, \omega + \Delta\bar{p}, p\omega = 10 + (-0.7778) + 2.2777 = 11.5$ хил. лева. Същото увеличение $\Delta\bar{p}$ може да се представи и като сума на общите приноси на всяка i -та стока с нейните три ефекта, или $\sum (\Delta p_i \omega_i + \Delta \omega_i p_i + h_i \Delta p_i \Delta \omega_i) = \sum (p_{i1} \omega_{i1} - p_{i0} \omega_{i0}) = \Delta\bar{p}$, откъдето $14.2222 + (-6.6667) + 2.5000 + 1.4444 = 11.5$ хил. лева (табл. 1, колона 10).

С получените абсолютни ефекти от адитивния факторен анализ се преминава в неговата относителна форма:

$$\frac{\Delta \bar{p}}{\bar{p}_0} = \frac{11.5}{53.5} = 0.2150, \quad \frac{\Delta \bar{p}, p}{\bar{p}_0} = \frac{10}{53.5} = 0.1869,$$

$$\frac{\Delta \bar{p}, \omega}{\bar{p}_0} = \frac{-0.7778}{53.5} = -0.0145 \text{ и } \frac{\Delta \bar{p}, p\omega}{\bar{p}_0} = \frac{2.2778}{53.5} = 0.0426.$$

Алгебричната сума на относителните ефекти е равна на относителния прираст на базисната средна $\frac{\Delta \bar{p}}{\bar{p}_0} = 0.2150$, тъй като $\frac{\Delta \bar{p}, p}{\bar{p}_0} + \frac{\Delta \bar{p}, \omega}{\bar{p}_0} + \frac{\Delta \bar{p}, p\omega}{\bar{p}_0} = 0.1869 + (-0.0145) + 0.0426 = 0.2150$. Интерпретацията на тези ефекти е, че нетните преобладаващи увеличения на цените на стоките са повлияли за голямото относително увеличение на средната базисна цена \bar{p}_0 с 18.69%, докато нетните преобладаващи намаления на относителните дялове са повлияли за много по-слабото намаление на \bar{p} с -1.45%. От своя страна, преобладаващите съвместни увеличения на цените на стоките и на техните относителни дялове са повлияли за относителното увеличение на \bar{p}_0 с 4.26%. Или общото процентно увеличение на средната базисна цена с $\frac{\Delta \bar{p}}{\bar{p}_0} = 21.50\%$ е равно на алгебричната сума на относителните (процентни) ефекти с $18.69\% + (-1.45\%) + 4.26\% = 21.50\%$.

2.2. Адитивен факторен анализ на средните цени от групирани данни на стоки с ефектите $\Delta \bar{p}, p < 0$, $\Delta \bar{p}, \omega > 0$ и $\Delta \bar{p}, p\omega < 0$

Следващият пример е за обратния случай на предходния в точка 2.1. Входните данни за него са представени в Приложение 2 и в табл. 2.

2. Адитивен факторен анализ на средните цени на стоките

Стоки	Относителни дялове		Ефекти от промени на цените			Ефекти от промени на относителните дялове			Съвместни ефекти	$p_{i1}\omega_{i1} - p_{i0}\omega_{i0}$
	ω_{i0}	ω_{i1}	Δp_i	ω_{im}	$\Delta p_i \omega_{im}$	$\Delta \omega_i$	p_{im}	$\Delta \omega_i p_{im}$		
í	1	2	3	4	5=3x4	6=2-1	7	8=6x7	9	10=5+8+9
1	0.2777	0.2000	-40	0.2000	-8.0000	-0.0777	40	-3.1111	-3.1111	-14.2222
2	0.1667	0.2500	+10	0.1667	1.6667	0.0833	50	4.1667	0.8333	6.6667
3	0.1667	0.2500	-40	0.1667	-6.6667	0.0833	50	4.1667	-	-2.5000
4	0.3889	0.3000	+10	0.3000	3.0000	-0.0889	50	-4.4444	-	-1.4444
Общо	1.0000	1.0000	X	X	-10	0	X	0.7778	-2.2778	-11.5

От данните в табл. 2 се установява, че средната цена през базисната година е $\bar{p}_0 = \frac{\sum p_{i0}q_{i0}}{Q_0} = \sum p_{i0}\omega_{i0} = 65.0$ хил. лв., а за отчетната година е $\bar{p}_1 = \frac{\sum p_{i1}q_{i1}}{Q_1} = \sum p_{i1}\omega_{i1} = 53.5$ хил. лева. Или средната цена на стоките е намаляла през отчетната година с $\Delta\bar{p} = \bar{p}_1 - \bar{p}_0 = 53.5 - 65.0 = -11.5$ хил. лева (табл. 2, колона 10). Това намаление се разпределя на отделните ефекти:

– ценовият $\Delta\bar{p},p$ само от промените на цените е намалението на средната базисна цена с $\Delta\bar{p},p = \sum \Delta p_i \omega_{imin} = -10$ хил. лева (табл. 2, колона 5). Същото намаление се образува от общия по-голям отрицателен ефект -14.6667 хил. лв. от преобладаващото влияние на намалените цени на първата и третата стока, и от общия по-малък положителен ефект 4.6667 хил. лв. от по-слабите увеличения на цените на втората и четвъртата стока. Приносите на отделните стоки за ефекта $\Delta\bar{p},p$ са отделните ценови ефекти за всяка стока в колона 5 на табл. 2.

– структурният ефект $\Delta\bar{p},\omega$ само от промените на относителните дялове на натуралните количества на стоките е минималното увеличение на \bar{p}_0 с $\Delta\bar{p},\omega = \sum \Delta \omega_i p_{imin} = 0.7778$ хил. лева (табл. 2, колона 8). Това минимално увеличение произлиза от общия малко по-голям положителен ефект 8.3333 хил. лв. от преобладаващото влияние на увеличенията на относителните дялове на втората и третата стока и от общия по-малък отрицателен ефект -7.5555 хил. лв. от намаленията на относителните дялове на първата и четвъртата стока. Приносите на отделните стоки за ефекта $\Delta\bar{p},\omega$ са отделните структурни ефекти за всяка стока в колона 8 на табл. 2.

– съвместният ефект $\Delta\bar{p},p\omega$ само от еднопосочните съвместни промени на цените на стоките и на относителните дялове на техните натурални количества е намалението на средната базисна цена с $\Delta\bar{p},p\omega = \sum h_i \Delta p_i \Delta \omega_i = -2.2778$ хил. лева (табл. 2, колона 9). Посоченото намаление на \bar{p}_0 се получава от по-големия по абсолютна стойност отрицателен съвместен ефект -3.1111 хил. лв. на първата стока и от по-малкия положителен съвместен ефект 0.8333 хил. лв. на втората стока.

Общото намаление на средната базисна цена с $\Delta\bar{p} = -11.5$ хил. лв се представя с алгебричната сума на трите ефекта $\Delta\bar{p},p + \Delta\bar{p},\omega + \Delta\bar{p},p\omega = -10 + 0.7778 + (-2.2777) = -11.5$ хил. лева. Това намаление на \bar{p}_0 може да се представи и като сума на общите приноси на всяка i -та стока с нейните три ефекта $\sum (\Delta p_i \omega_i + \Delta \omega_i p_i + h_i \Delta p_i \Delta \omega_i) = \sum (p_{i1} \omega_{i1} - p_{i0} \omega_{i0}) = \Delta\bar{p}$, откъдето $-14.2222 + 6.6667 + (-2.5000) + (-1.4444) = -11.5$ хил. лева (табл. 2, колона 10).

С решения пример, който е обратен на първия пример в т. 2.1, се потвърждава правилото за взаимнообратимостта на ефектите от случаите, които са с едни и същи данни, но са с разменени места на базисната и отчетната година. От решението на втория пример се получават същите ефекти по абсолютна стойност както от първия пример, но са с обратни алгебрични знаци. От решението на абсолютната форма на адитивния факторен анализ се преминава в неговата относителна форма:

$$\frac{\Delta \bar{p}}{\bar{p}_0} = \frac{-11.5}{65.0} = -0.1769, \quad \frac{\Delta \bar{p}, p}{\bar{p}_0} = \frac{-10}{65.0} = -0.1539,$$

$$\frac{\Delta \bar{p}, \omega}{\bar{p}_0} = \frac{0.7778}{65.0} = 0.0120 \text{ и } \frac{\Delta \bar{p}, p\omega}{\bar{p}_0} = \frac{-2.2778}{65.0} = -0.0350.$$

Алгебричната сума на относителните ефекти е равна на относителното намаление на базисната средна $\frac{\Delta \bar{p}}{\bar{p}_0} = -0.1769$, защото $\frac{\Delta \bar{p}, p}{\bar{p}_0} + \frac{\Delta \bar{p}, \omega}{\bar{p}_0} + \frac{\Delta \bar{p}, p\omega}{\bar{p}_0} = -0.1539 + 0.0120 + (-0.0350) = -0.1769$. Интерпретацията на получените ефекти е, че нетните преобладаващи намаления на цените на стоките са повлияли за голямото относително намаление на средната базисна цена \bar{p}_0 с -15.39% , докато нетните преобладаващи увеличения на относителните дялове на натуралните количества на стоките са повлияли за много по-слабото относително увеличение на \bar{p}_0 с 1.20% . Освен влиянията на тези факторни промени преобладаващите съвместни намаления на цените на стоките и на техните относителни дялове са повлияли за относителното намаление на \bar{p}_0 с още -3.50% . Или общото относително (процентно) намаление на средната базисна цена с $\frac{\Delta \bar{p}}{\bar{p}_0} = -17.69\%$ е равно на алгебричната сума на относителните (процентни) ефекти $-15.39\% + 1.20\% + (-3.50\%) = -17.69\%$.

2.3. и 2.4. Адитивен факторен анализ на средните цени от групирани данни на стоки за случая с ефектите $\Delta \bar{p}, p > 0$, $\Delta \bar{p}, \omega < 0$ и $\Delta \bar{p}, p\omega < 0$ и за случая с ефектите $\Delta \bar{p}, p < 0$, $\Delta \bar{p}, \omega > 0$ и $\Delta \bar{p}, p\omega > 0$

Следващите два случая са третият и четвъртият от схемата за осемте случая на адитивния факторен анализ в точка 1. Според възприетата номерация в точка 2 двата случая трябва да бъдат с номера 2.3 и 2.4. С цел да бъдат по-кратки предходните публикации на автора, за тези случаи не са решавани примери с агрегирани данни на стоките, поради което и тук няма да бъдат решени с групирани данни (Христов, 2016а, 2016б). Ако читателят желае, може да състави такива примери с агрегирани и

групирани данни и да ги реши безпроблемно с дискретната нечетна функция на математическия сигнал както предходните два примера в точки 2.1. и 2.2.

2.5. Адитивен факторен анализ на средните цени от групирани данни на стоки с ефектите $\Delta \bar{p}, p > 0$, $\Delta \bar{p}, \omega > 0$ и $\Delta \bar{p}, p\omega < 0$

Входните данни за примера на този случай са поместени в Приложение 3 и в табл.

3.

3. Адитивен факторен анализ на средните цени на стоките

Стоки	Относителни дялове		Ефекти от промени на цените			Ефекти от промени на относителните дялове			Съвместни ефекти	$p_{i1}\omega_{i1} - p_{i0}\omega_{i0}$
	ω_{i0}	ω_{i1}	Δp_i	ω_{im}	$\Delta p_i \omega_{im}$	$\Delta \omega_i$	p_{im}	$\Delta \omega_i p_{im}$		
í	1	2	3	4	5=3x4	6=2-1	7	8=6x7	9	10=5+8+9
1	0.3103	0.3572	10	0.3103	3.1030	0.0469	90	4.2210	0.4690	7.7930
2	0.2759	0.3035	10	0.2759	2.7590	0.0276	80	2.2080	0.2760	5.2430
3	0.1724	0.1250	-20	0.1250	-2.5000	-0.0474	60	-2.8440	-0.9480	-6.2920
4	0.0690	0.0357	-10	0.0357	-0.3570	0.0333	30	-0.9990	-0.3330	-1.6890
5	0.1034	0.0714	40	0.0714	2.8560	0.0320	20	-0.6400	-	2.2160
6	0.0690	0.1072	-30	0.0690	-2.0700	0.0382	50	1.9100	-	-0.1600
Общо	1.0000	1.0000	X	X	8.7180	0	X	8.3390	0.7450	15.2520
					-4.9270			-4.4830	-1.2810	-8.1410
					3.7910			3.8560	-0.5360	7.1110

Според посочените данни в табл. 3 средните цени са $\bar{p}_0 = \frac{\sum p_{i0}q_{i0}}{Q_0} = \sum p_{i0}\omega_{i0} = 74.138$ хил. лв. за базисната година и $\bar{p}_1 = \frac{\sum p_{i1}q_{i1}}{Q_1} = \sum p_{i1}\omega_{i1} = 81.250$ хил. лв. за отчетната година. Увеличението на средната базисна цена е с $\Delta \bar{p} = \bar{p}_1 - \bar{p}_0 = 81.250 - 74.138 = 7.112$ хил. лева. Получената стойност за $\Delta \bar{p} = \bar{p}_1 = 7.1110$ хил. лв. в колона 10 на табл. 3 е поради закръгляния на \bar{p}_0 и \bar{p}_1 до третия десетичен знак.

Увеличението $\Delta \bar{p} = 7.1110$ хил. лв. се разпределя на следните ефекти:

– ценовият $\Delta \bar{p}, p$ само от промените на цените на стоките е увеличение на средната базисна цена с $\Delta \bar{p}, p = \sum \Delta p_i \omega_{imin} = 3.7910$ хил. лева (табл. 3, колона 5). Той се образува от общия по-голям положителен ефект 8.7180 хил. лв. от преобладаващото влияние на увеличените цени на първата, втората и петата стока и от общия по-малък

отрицателен ефект -4.9270 хил. лв. от по-слабите намаления на цените на третата, четвъртата и шестата стока. Приносите на отделните стоки за ефекта $\Delta\bar{p}$, p са отделните ценови ефекти за всяка стока в колона 5 на табл. 3.

– структурният ефект $\Delta\bar{p}$, ω само от промените на относителните дялове на натуралните количества на стоките възлиза на $\Delta\bar{p}$, $\omega = \sum \Delta\omega_i p_{imin} = 3.8560$ хил. лв. и представлява увеличение на \bar{p}_0 (табл. 3, колона 8). Той е незначително по-голям от ценовия ефект $\Delta\bar{p}$, $p = 3.7910$ хил. лева. Същият структурен ефект произлиза от по-големия положителен ефект 8.3390 хил. лв. от преобладаващото влияние на увеличението на относителните дялове на първата, втората и шестата стока и от общия по-малък отрицателен ефект -4.4830 хил. лв. от намаленията на относителните дялове на третата, четвъртата и петата стока. Приносите на отделните стоки за ефекта $\Delta\bar{p}$, ω са отделните структурни ефекти за всяка стока в колона 8 на табл. 3.

– съвместният ефект $\Delta\bar{p}$, $p\omega$ само от еднопосочните съвместни промени на цените на стоките и на относителните дялове на техните натурални количества е минималното намаление на \bar{p}_0 с $\Delta\bar{p}$, $p\omega = \sum h_i \Delta p_i \Delta \omega_i = -0.5360$ хил. лева (табл. 3, колона 9). Този ефект се получава от общия по-голям отрицателен съвместен ефект -1.2810 хил. лв. от третата и четвъртата стока и от общия по-малък положителен съвместен ефект 0.7450 хил. лв. от първата и втората стока (табл. 3, колона 9). Последните две стоки (петата и шестата) са само с разнопосочни факторни промени, от които няма съвместни ефекти. Или общото увеличение на средната базисна цена с $\Delta\bar{p} = 7.1110$ хил. лв. се представя с алгебричната сума на трите ефекта $\Delta\bar{p}$, $p + \Delta\bar{p}$, $\omega + \Delta\bar{p}$, $p\omega = 3.7910 + 3.8560 + (-0.5360) = 7.1110$ хил. лева. Същото увеличение на \bar{p}_0 може да се представи и със сумата на общите приноси на всяка i -та стока с нейните три ефекта, или $\sum(\Delta p_i \omega_i + \Delta \omega_i p_i + h_i \Delta p_i \Delta \omega_i) = \sum(p_{i1} \omega_{i1} - p_{i0} \omega_{i0}) = \Delta\bar{p}$, откъдето $7.7930 + 5.2430 + (-6.2920) + (-1.6890) + 2.2160 + (-0.1600) = 15.2520 - 8.1410 = 7.1110$ хил. лева (табл. 3, колона 10).

От абсолютната форма на адитивния факторен анализ се преминава в неговата относителна форма:

$$\frac{\Delta\bar{p}}{\bar{p}_0} = \frac{7.111}{74.138} = 0.0959, \quad \frac{\Delta\bar{p}, p}{\bar{p}_0} = \frac{3.791}{74.138} = 0.0511,$$

$$\frac{\Delta \bar{p}, \omega}{\bar{p}_0} = \frac{3.856}{74.138} = 0.0520 \text{ и } \frac{\Delta \bar{p}, p\omega}{\bar{p}_0} = \frac{-0.536}{74.138} = -0.0072.$$

Алгебричната сума на тези относителни ефекти е равна на относителния прираст на базисната средна цена $\frac{\Delta \bar{p}}{\bar{p}_0} = 0.0959$, защото $0.0511 + 0.0520 + (-0.0072) = 0.0959$. Интерпретацията на относителните ефекти е, че нетните преобладаващи увеличения на цените на стоките са повлияли за относителното увеличение на средната базисна цена \bar{p}_0 с 5.11%, докато нетните преобладаващи увеличения на относителните дялове са повлияли за малко по-голямо относително увеличение на \bar{p}_0 с 5.20%. Според съвместния ефект преобладаващите съвместни намаления на цените и на техните относителни дялове са повлияли за много малкото относително намаление на \bar{p}_0 с -0.72%. Или общото процентно увеличение на средната базисна цена с $\frac{\Delta \bar{p}}{\bar{p}_0} = 9.59\%$ е равно на алгебричната сума на относителните (процентни) ефекти $5.11\% + 5.20\% + (-0.72\%) = 9.59\%$.

2.6. Адитивен факторен анализ на средните цени от групирани данни на стоки с ефектите $\Delta \bar{p}, p < 0$, $\Delta \bar{p}, \omega < 0$ и $\Delta \bar{p}, p\omega > 0$

Примерът за този случай е обратен на решения пример в предходната точка 2.5. Входните данни за него са поместени в Приложение 4 и в табл. 4.

4. Адитивен факторен анализ на средните цени на стоките

Стоки	Относителни дялове		Ефекти от промени на цените			Ефекти от промени на относителните дялове			Съвместни ефекти	$p_{i1}\omega_{i1} - p_{i0}\omega_{i0}$
	ω_{i0}	ω_{i1}	Δp_i	ω_{im}	$\Delta p_i \omega_{im}$	$\Delta \omega_i$	p_{im}	$\Delta \omega_i p_{im}$		
i	1	2	3	4	5=3x4	6=2-1	7	8=6x7	9	10=5+8+9
1	0.3572	0.3103	-10	0.3103	-3.1030	-0.0469	90	-4.2210	-0.4690	-7.7930
2	0.3035	0.2759	-10	0.2759	-2.7590	-0.0276	80	-2.2080	-0.2760	-5.2430
3	0.1250	0.1724	20	0.1250	2.5000	0.0474	60	2.8440	0.9480	6.2920
4	0.0357	0.0690	10	0.0357	0.3570	0.0333	30	0.9990	0.3330	1.6890
5	0.0714	0.1034	-40	0.0714	-2.8560	0.0320	20	0.6400	-	-2.2160
6	0.1072	0.0690	30	0.0690	2.0700	-0.0382	50	-1.9100	-	0.1600
Общо	1.0000	1.0000	X	X	4.9270	0	X	4.4830	1.2810	8.1410
					-8.7180			-8.3390	-0.7450	-15.2520
					-3.7910			-3.8560	0.5360	-7.1110

От данните в табл. 4 двете средни цени са $\bar{p}_0 = \frac{\sum p_{i0}q_{i0}}{Q_0} = \sum p_{i0}\omega_{i0} = 81.250$ хил. лв. за базисната година и $\bar{p}_1 = \frac{\sum p_{i1}q_{i1}}{Q_1} = \sum p_{i1}\omega_{i1} = 74.138$ хил. лв. за отчетната година. Или за отчетната година средната базисна цена \bar{p}_0 е намаляла с $\Delta \bar{p} = \bar{p}_1 - \bar{p}_0 = 74.138 - 81.250 = -7.112$ хил. лева. По-точната стойност на $\Delta \bar{p} = -7.1110$ хил. лв. в колона 10 на табл. 4 се дължи както в предходния пример на закръгления на средните

цени на \bar{p}_0 и \bar{p}_1 до третия десетичен знак. Намалението $\Delta\bar{p} = -7.1110$ се състои от следните ефекти:

– ценовият $\Delta\bar{p}, p$ само от промените на цените на стоките е намалението на средната базисна цена с $\Delta\bar{p}, p = \sum \Delta p_i \omega_{imin} = -3.7910$ хил. лева (табл. 4, колона 5). То се образува от общия по-голям отрицателен ефект -8.7180 хил. лв. от преобладаващото влияние на намалелите цени на първата, втората и петата стока и от общия по-малък положителен ефект 4.9270 хил. лв. от по-слабите увеличения на цените на третата, четвъртата и шестата стока. Приносите на отделните стоки за ефекта $\Delta\bar{p}, p$ са отделните ценови ефекти за всяка стока в колона 5 на табл. 4.

– структурният ефект $\Delta\bar{p}, \omega$ само от промените на относителните дялове на натуралните количества на стоките е $\Delta\bar{p}, \omega = \sum \Delta \omega_i p_{imin} = -3.8560$ хил. лв. увеличение на \bar{p}_0 (табл. 4, колона 8). Той е незначително по-голям от ценовия ефект $\Delta\bar{p}, p = -3.7910$ хил. лева. Същият структурен ефект произлиза от по-големия по абсолютна стойност отрицателен ефект -8.3390 хил. лв. от преобладаващото влияние на намаленията на относителните дялове на първата, втората и шестата стока и от общия по-малък положителен ефект 4.4830 хил. лв. от увеличенията на относителните дялове на третата, четвъртата и петата стока. Приносите на отделните стоки за ефекта $\Delta\bar{p}, \omega$ са структурните ефекти за всяка стока в колона 8 на табл. 4.

– съвместният ефект $\Delta\bar{p}, p\omega$ от еднопосочните факторни промени на цените на стоките и на относителните дялове на техните натурални количества е минималното увеличение на \bar{p}_0 с $\Delta\bar{p}, p\omega = \sum h_i \Delta p_i \Delta \omega_i = 0.5360$ хил. лева (табл. 4, колона 9). Този ефект се получава от общия по-голям положителен съвместен ефект 1.2810 хил. лв. от третата и четвъртата стока и от общия по-малък отрицателен съвместен ефект -0.7450 хил. лв. от първата и втората стока (табл. 4, колона 9). Последните две стоки (петата и шестата) са само с разнопосочни факторни промени, от които няма съвместни ефекти. Общото намаление на средната базисна цена с $\Delta\bar{p} = -7.1110$ хил. лв. се представя с алгебричната сума на трите ефекта $\Delta\bar{p}, p + \Delta\bar{p}, \omega + \Delta\bar{p}, p\omega = -3.7910 + (-3.8560) + 0.5360 = -7.1110$ хил. лева. Това намаление на \bar{p}_0 може да се представи и с алгебричната сума на общите приноси на всяка i -та стока в колона 10 на табл. 4:

$$\begin{aligned} \sum (\Delta p, \omega_i + \Delta \omega, p_i + h_i \Delta p_i \Delta \omega_i) &= \sum (p_{i1} \omega_{i1} - p_{i0} \omega_{i0}) = \Delta\bar{p}, \text{откъдето } -7.7930 + \\ &+ (-5.2430) + 6.2920 + 1.6890 + (-2.2160) + 0.1600 = -15.2520 + 8.1410 = \\ &-7.1110 \text{ хил. лева.} \end{aligned}$$

Решеният пример е взаимнообратим с предходния пример в т. 2.5, поради което получените ефекти от двата примера са равни по абсолютна стойност, но са с обратни алгебрични знаци.

От абсолютната форма на адитивния факторен анализ се преминава в неговата относителна форма:

$$\frac{\Delta \bar{p}}{\bar{p}_0} = \frac{-7.111}{81.250} = -0.0875, \quad \frac{\Delta \bar{p}, p}{\bar{p}_0} = \frac{-3.791}{81.250} = -0.0467,$$

$$\frac{\Delta \bar{p}, \omega}{\bar{p}_0} = \frac{-3.856}{81.250} = -0.0475 \text{ и } \frac{\Delta \bar{p}, p\omega}{\bar{p}_0} = \frac{0.536}{81.250} = 0.0066.$$

Алгебричната сума на тези относителни ефекти е равна на относителното намаление на базисната средна цена $\frac{\Delta \bar{p}}{\bar{p}_0} = -0.0875$, защото $-0.0467 + (-0.0475) + 0.0066 = -0.0942 + 0.0066 = -0.0876$. Разликата на този резултат с относителното намаление на средната базисна цена $\frac{\Delta \bar{p}}{\bar{p}_0} = -0.0875$ се дължи на закръгления на $\Delta \bar{p}, p$ и $\Delta \bar{p}, \omega$.

Интерпретацията на относителните ефекти е, че нетните преобладаващи намаления на цените на стоките са повлияли за относителното намаление на средната базисна цена \bar{p}_0 с -4.67% , докато нетните преобладаващи намаления на относителните дялове на натуралните количества на стоките са повлияли за малко по-голямото относително намаление на \bar{p}_0 с -4.75% . Само преобладаващите съвместни увеличения на цените на стоките и на техните относителни дялове са повлияли за много малкото относително увеличение на \bar{p}_0 с 0.66% . По този начин общото процентно намаление на средната базисна цена с $\frac{\Delta \bar{p}}{\bar{p}_0} = -8.75\%$ е равно на алгебричната сума на относителните (процентни) ефекти $-4.67\% + (-4.75\%) + 0.66\% = -8.76\%$.

2.7. Адитивен факторен анализ на средните цени от групирани данни на стоки с ефектите $\Delta \bar{p}, p > 0$, $\Delta \bar{p}, \omega > 0$ и $\Delta \bar{p}, p\omega > 0$

Последните два случая от схемата в точка 2 са с еднопосочните факторни промени от **преобладаващи** едновременни увеличения или намаления на цените и на относителните дялове на стоките. Двата примера за тези случаи са съставени с четири стоки, продукцията на всяка от които се анализира с един от двата вида еднопосочни факторни промени. Както е отбелязано в методиката за адитивния факторен анализ на средните равнища (цени) в предходната точка 1, и тук е необходимо да се отбележи

относителните дялове на стоките чрез дискретната функция на математическия сигнум. Според него промяната на всеки от двата фактора трябва да се претегля (умножава) **винаги с по-малкото равнище** на другия фактор от базисната или от отчетната година. При адитивния факторен анализ с цените p_i и с техните натурални количества q_i са възможни едновременни увеличения или намаления и на двата фактора при някои наблюдавани стоки. Промените само на относителните дялове ω_i на една част от стоките са **винаги** разнопосочни. Една част, дори и само един от относителните дялове да се е увеличил, някои, дори и само един от относителните дялове е намалял. Тогава, **ако** цените p_i на всички наблюдавани стоки са се увеличили и те се претеглят **само** с относителните дялове ω_{i0} от базисната година или **само** с ω_{i1} от отчетната година, **една част от стоките ще бъде с неверни ефекти**. По същия начин, **ако** промените на относителните дялове ω_i се умножат **само** с цените p_{i0} от базисната година или само с p_{i1} от отчетната година, **една част от стоките ще бъде също с неверни ефекти**.

Първият от двата примера за еднопосочните факторни промени е с трите положителни ефекта $\Delta \bar{p}, p > 0$, $\Delta \bar{p}, \omega > 0$ и $\Delta \bar{p}, p\omega > 0$. Входните данни за този пример са представени в Приложение 5 и в табл. 5.

5. Адитивен факторен анализ на средните цени на стоките

Стоки	Относителни дялове		Ефекти от промени на цените			Ефекти от промени на относителните дялове			Съвместни ефекти	$p_{i1}\omega_{i1} - p_{i0}\omega_{i0}$
	ω_{i0}	ω_{i1}	Δp_i	ω_{im}	$\Delta p_i \omega_{im}$	$\Delta \omega_i$	p_{im}	$\Delta \omega_i p_{im}$		
i	1	2	3	4	5=3x4	6=2-1	7	8=6x7	9	10=5+8+9
1	0.2222	0.3000	+40	0.2222	8.8889	+0.0778	40	+3.1111	+3.1111	+15.1111
2	0.2778	0.2000	-20	0.2000	-4.0000	-0.0778	40	-3.1111	-1.5556	-8.6667
3	0.2778	0.2000	+20	0.2000	+4.0000	-0.0778	40	-3.1111	-	+0.8889
4	0.2222	0.3000	-10	0.2222	-2.2222	+0.0778	50	+3.8889	-	+1.6667
Общо	1.0000	1.0000	X	X	+6.6667	0	X	+0.7778	+1.5555	+9.0000

Според данните в табл. 5 претеглените средни цени са $\bar{p}_0 = \frac{\sum p_{i0}q_{i0}}{Q_0} = \sum p_{i0}\omega_{i0} = 50$ хил. лв. за базисната година и $\bar{p}_1 = \frac{\sum p_{i1}q_{i1}}{Q_1} = \sum p_{i1}\omega_{i1} = 59$ хил. лв. за отчетната година. Или средната цена \bar{p}_0 се е увеличила с $\Delta \bar{p} = \bar{p}_1 - \bar{p}_0 = 59 - 50 = 9$ хил. лева (табл. 5, колона 10). Същото увеличение е сумата на следните положителни ефекти:

– ценовият $\Delta \bar{p}, p$ само от промените на цените на стоките е увеличението на средната базисна цена с $\Delta \bar{p}, p = \sum \Delta p_i \omega_{imin} = 6.6667$ хил. лева (табл. 5, колона 5). Този ефект се образува от общия по-голям положителен ефект 12.8889 хил. лв. от преобладаващото влияние на увеличените цени на първата и третата стока и от общия по-малък отрицателен ефект – 6.2222 хил. лв. от по-слабите намаления на цените на

втората и четвъртата стоки. Приносите на отделните стоки за ефекта $\Delta\bar{p}, p$ са отделните ценови ефекти за всяка стока в колона 5 на табл. 5.

– структурният ефект $\Delta\bar{p}, \omega$ само от промените на относителните дялове на натуралните количества на стоките е много по-малкото увеличение на \bar{p}_0 с $\Delta\bar{p}, \omega = \sum \Delta\omega_i p_{imin} = 0.7778$ хил. лева (табл. 5, колона 8). Това увеличение се получава от малко по-големия общ положителен ефект 7.0000 хил. лв. от преобладаващото влияние на увеличените относителни дялове на първата и четвъртата стока и от общия по-малък отрицателен ефект -6.2222 хил. лв. от намаленията на относителните дялове на втората и третата стока. Приносите на отделните стоки за ефекта $\Delta\bar{p}, \omega$ са отделните структурни ефекти за всяка стока в колона 8 на табл. 5.

– съвместният ефект $\Delta\bar{p}, p\omega$ само от еднопосочните съвместни промени на цените на стоките и на относителните дялове на техните натурални количества е увеличението на \bar{p}_0 с $\Delta\bar{p}, p\omega = \sum h_i \Delta p_i \Delta \omega_i = 1.5555$ хил. лева (табл. 5, колона 9). Това увеличение се получава от по-големия положителен съвместен ефект 3.1111 хил. лв. на първата стока и от по-малкия отрицателен съвместен ефект -1.5555 хил. лв. на втората стока. Другите две стоки (третата и четвъртата) са само с разнопосочни факторни промени, от които няма съвместни ефекти.

Общото увеличение на средната базисна цена \bar{p}_0 с $\Delta\bar{p}=9$ хил. лв. е равно на сумата на трите положителни ефекта: $\Delta\bar{p}, p + \Delta\bar{p}, \omega + \Delta\bar{p}, p\omega = 6.6667 + 0.7778 + 1.5555 = 9.0000$ хил. лева. Това увеличение може да се представи и със сумата на общите приноси с трите ефекта на всяка i -та стока, или $\sum(\Delta p, \omega_i + \Delta \omega, p_i + h_i \Delta p_i \Delta \omega_i) = \sum(p_{i1}\omega_{i1} - p_{i0}\omega_{i0}) = \Delta\bar{p}$, откъдето $15.1111 - 8.6667 + 0.8889 + 1.6667 = 17.6667 - 8.6667 = 9.0000$ хил. лева (табл. 5, колона 10).

От адитивната форма се преминава в относителната форма на адитивния факторен анализ:

$$\frac{\Delta\bar{p}}{\bar{p}_0} = \frac{9}{50} = 0.1800, \quad \frac{\Delta\bar{p}, p}{\bar{p}_0} = \frac{6.6667}{50} = 0.1333,$$

$$\frac{\Delta\bar{p}, \omega}{\bar{p}_0} = \frac{0.7778}{50} = 0.0156 \text{ и } \frac{\Delta\bar{p}, p\omega}{\bar{p}_0} = \frac{1.5555}{50} = 0.0311.$$

Абсолютната сума на относителните ефекти е равна на относителния прираст на базисната средна цена $\frac{\Delta\bar{p}}{\bar{p}_0} = 0.1800$, тъй като $0.1333 + 0.0156 + 0.0311 = 0.1800$.

Интерпретацията на получените ефекти е, че нетните преобладаващи увеличения на

цените на стоките са повлияли за голямото относително увеличение на средната базисна цена \bar{p}_0 с 13.33%. Другото нетно увеличение на \bar{p}_0 е много по-малко - само с 1.56% от преобладаващите увеличения на относителните дялове на натуралните количества на стоките. Третото увеличение на \bar{p}_0 е с 3.11%, което произлиза само от преобладаващите съвместни увеличения на цените и на относителните дялове на първата стока и съвместните намаления на втората стока. Или общото процентно увеличение на \bar{p}_0 с $\frac{\Delta \bar{p}}{\bar{p}_0} = 18.00\%$ е равно на сумата на относителните (процентни) положителни ефекти $13.33\% + 1.56\% + 3.11\% = 18.00\%$. Подобно на пропорционалното разпределение на съвместния ефект от адитивния факторен анализ на продукцията от разнородните съвкупности на стоките в моята предходна статия и тук съвместният ефект от адитивния анализ на средните цени може да се разпредели пропорционално с два брутни факторни индекса (Христов, 2017). Те се определят с известното квадратно уравнение, което ще бъде изведено в следващата публикация за индексния факторен анализ на средните цени.

2.8. Адитивен факторен анализ на средните цени от групирани данни на стоки с ефектите $\Delta \bar{p}, p < 0$, $\Delta \bar{p}, \omega < 0$ и $\Delta \bar{p}, p\omega < 0$

Този случай е последният от схемата в точка 2 и е обратен на предходния случай с трите положителни ефекта в точка 2.7. Началните данни за неговия пример са изложени в Приложение 6 и табл. 6.

6. Адитивен факторен анализ на средните цени на стоките

Стоки	Относителни дялове		Ефекти от промени на цените			Ефекти от промени на относителните дялове			Съвместни ефекти	$p_{i1}\omega_{i1} - p_{i0}\omega_{i0}$
	ω_{i0}	ω_{i1}	Δp_i	ω_{im}	$\Delta p_i \omega_{im}$	$\Delta \omega_i$	p_{im}	$\Delta \omega_i p_{im}$	$\Delta p_i \Delta \omega_i$	
í	1	2	3	4	5=3x4	6=2-1	7	8=6x7	9	10=5+8+9
1	0.3000	0.2222	-40	0.2222	-8.8889	-0.0778	40	-3.1111	-3.1111	-15.1111
2	0.2000	0.2778	+20	0.2000	+4.0000	+0.0778	40	+3.1111	+1.5556	8.6667
3	0.2000	0.2778	-20	0.2000	-4.0000	+0.0778	40	+3.1111	-	-0.8889
4	0.3000	0.2222	+10	0.2222	+2.2222	-0.0778	50	-3.8889	-	-1.6667
Общо	1.0000	1.0000	X	X	-6.6667	0	X	-0.7778	-1.5555	-9.0000

От данните в табл. 6 средните претеглени цени за базисната и отчетната година са $\bar{p}_0 = \frac{\sum p_{i0}q_{i0}}{Q_0} = \sum p_{i0}\omega_{i0} = 59$ хил. лв. и $\bar{p}_1 = \frac{\sum p_{i1}q_{i1}}{Q_1} = \sum p_{i1}\omega_{i1} = 50$ хил. лева. Според

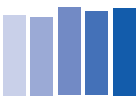
тях намалението на средната базисна цена \bar{p}_0 е с $\Delta\bar{p} = \bar{p}_1 - \bar{p}_0 = 50 - 59 = -9$ хил. лева (табл. 6, колона 10). Същото намаление е сума на трите отрицателни ефекта:

– ценовият $\Delta\bar{p}, p$ само от промените на цените на стоките възлиза на $\Delta\bar{p}, p = \sum \Delta p_i \omega_{imin} = -6.6667$ хил. лв. намаление на \bar{p}_0 (табл. 6, колона 5). Това намаление се образува от общия по-голям отрицателен ефект -12.8889 хил. лв. от преобладаващото влияние на намалелите цени на първата и третата стока и от общия по-малък положителен ефект 6.2222 хил. лв. от по-слабите увеличения на цените на втората и четвъртата стока. Приносите на отделните стоки за ефекта $\Delta\bar{p}, p$ са отделните ценови ефекти за всяка стока в колона 5 на табл. 6.

– структурният ефект $\Delta\bar{p}, \omega$ само от промените на относителните дялове на натуралните количества на стоките е много по-малкото намаление на \bar{p}_0 с $\Delta\bar{p}, \omega = \sum \Delta \omega_i p_{imin} = -0.7778$ хил. лева (табл. 6, колона 8). Същото намаление се получава от малко по-големия общ отрицателен ефект -7 хил. лв. от преобладаващото влияние на намалелите относителни дялове на първата и четвъртата стока и от общия по-малък положителен ефект 6.2222 хил. лв. от увеличенията на относителните дялове на втората и третата стока. Приносите на отделните стоки за ефекта $\Delta\bar{p}, \omega$ са отделните структурни ефекти за всяка стока в колона 8 на табл. 6.

– съвместният ефект $\Delta\bar{p}, p\omega$ само от еднопосочните съвместни промени на цените на стоките и на относителните дялове на техните натурални количества е намалението на \bar{p}_0 с $\Delta\bar{p}, p\omega = \sum h_i \Delta p_i \Delta \omega_i = -1.5555$ хил. лева (табл. 6, колона 9). Това намаление се получава от по-големия отрицателен съвместен ефект -3.1111 хил. лв. за първата стока и от по-малкия положителен съвместен ефект 1.5555 хил. лв. за втората стока. Другите две стоки (третата и четвъртата) са само с разнопосочни факторни промени, от които няма съвместни ефекти.

Общото намаление на средната базисна цена \bar{p}_0 с $\Delta\bar{p} = -9$ хил. лв. е равно на сумата на трите отрицателни ефекта: $\Delta\bar{p}, p + \Delta\bar{p}, \omega + \Delta\bar{p}, p\omega = -6.6667 + (-0.7778) + (-1.5555) = -9.0000$ хил. лева. Това намаление на \bar{p}_0 може да се представи и със сумата на общите приноси с трите ефекта на всяка i -та стока, или $\sum (\Delta p, \omega_i + \Delta \omega, p_i + h_i \Delta p_i \Delta \omega_i) = \sum (p_{i1} \omega_{i1} - p_{i0} \omega_{i0}) = \Delta\bar{p}$, откъдето $-15.1111 + 8.6667 + (-0.8889) + (-1.6667) = -17.6667 + 8.6667 = -9.0000$ хил. лева (табл. 6, колона 10).



Тъй като решеният пример е обратен на предходния в точка 2.7, с неговото решение се потвърждава правилото за двата взаимнообратими примера. Ефектите от тези примери са равни по абсолютна стойност, но са с различни алгебрични знаци.

От адитивната форма на анализа се преминава в неговата относителна форма:

$$\frac{\Delta \bar{p}}{\bar{p}_0} = \frac{9}{59} = -0.1526, \quad \frac{\Delta \bar{p}, p}{\bar{p}_0} = \frac{-6.6667}{59} = -0.1130,$$

$$\frac{\Delta \bar{p}, \omega}{\bar{p}_0} = \frac{-0.7778}{59} = -0.0132 \text{ и } \frac{\Delta \bar{p}, p\omega}{\bar{p}_0} = \frac{-1.5555}{59} = -0.0264.$$

Сумата на отрицателните ефекти е равна на относителното намаление на средната базисна цена с $\frac{\Delta \bar{p}}{\bar{p}_0} = -0.1526$, защото $-0.1130 + (-0.0132) + (-0.0264) = -0.1526$.

Интерпретацията на тези ефекти е, че нетните преобладаващи намаления на цените са повлияли за най-голямото относително намаление на средната базисна цена \bar{p}_0 с -11.30% . Следващото много по-малко нетно намаление на \bar{p}_0 е само с -1.32% от преобладаващите намаления на относителните дялове на стоките. Третото намаление на \bar{p}_0 е с -2.64% само от преобладаващите съвместни намаления на цените и на относителните дялове на стоките с такива промени. Или общото процентно намаление на \bar{p}_0 с $\frac{\Delta \bar{p}}{\bar{p}_0} = -15.26\%$ е равно на сумата на относителните (процентни) отрицателни ефекти $-11.30\% + (-1.32\%) + (-2.64\%) = -15.26\%$. Подобно на предходния пример с трите положителни ефекта и тук отрицателният съвместен ефект може да се разпредели пропорционално между двата нетни отрицателни ефекта.

3. Критика на методите за адитивен факторен анализ на средните равнища в икономическото образование и обществените науки у нас

Еднозначните решения на новата методика за адитивен факторен анализ на средните цени могат да се сравнят с решенията от подобни методики на дългогодишни и водещи специалисти по адитивен и индексен факторен анализ в катедра „Статистика и иконометрия“ на УНСС - София, която е водеща по статистика в икономическото образование у нас. Същите методи от тази катедра се прилагат и в други университети и колежи, както се прилагат и за анализи в някои обществени науки по икономика, демография и социология в БАН. По-конкретно, имам предвид методите на проф. Кирил Гатев и проф. Димитър Аркадиев, на доц. Тодор Къналиев и доц. Андреана Стойкова-Къналиева, както и на покойните професори Здравко Сугарев, Венец Цонев и Божидар Русев. Нарочно цитирам проф. З. Сугарев и проф. Б. Русев, защото са били

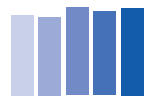
преподаватели и по демографска статистика, която е пример за прякото приложение на теорията на вероятностите в демографските анализи. Единствените верни решения на адитивния факторен анализ от икономическото образование са за изменението на **продукцията Р от еднородни съвкупности на стоките с агрегирани данни за средните цени \bar{p} и общите натурални количества на стоките Q**. Това са трите известни много отдавна решения от съветски източник (Югенбург, 1955). За разглеждания адитивен факторен анализ на продукцията от еднородните съвкупности на стоките трите решения са:

$$\Delta P = P_1 - P_0 = (\bar{p}_1 - \bar{p}_0)Q_0 + (Q_1 - Q_0)\bar{p}_0 + (\bar{p}_1 - \bar{p}_0)(Q_1 - Q_0),$$

$$\Delta P = P_1 - P_0 = (\bar{p}_1 - \bar{p}_0)Q_0 + (Q_1 - Q_0)\bar{p}_1 \text{ и}$$

$$\Delta P = P_1 - P_0 = (\bar{p}_1 - \bar{p}_0)Q_1 + (Q_1 - Q_0)\bar{p}_0 .$$

При $\Delta\bar{p}_1 < 1$ и $\Delta Q < 1$ четвъртото възможно решение е $\Delta P = P_1 - P_0 = (\bar{p}_1 - \bar{p}_0)Q_1 + (Q_1 - Q_0)\bar{p}_1 - (\bar{p}_1 - \bar{p}_0)(Q_1 - Q_0)$. То е изведено от мен през 1978 г. (Христов, 1978). За съжаление, с изключение на първото решение, най-разпространените решения на адитивния факторен анализ от икономическото образование у нас за изменението на продукцията от еднородни съвкупности на стоките с агрегираните данни за \bar{p} и Q са **неверни**. Ще посоча само двама автори - проф. Кирил Гатев, който е първоизточник на тези решения, и един от последните съвременни разпространители на същите решения - проф. Димитър Аркадиев (Гатев, 1995; Аркадиев, 2017). Основната причина за техните неверни решения е, че извеждат адитивния факторен анализ от предварителен **неверен** индексен факторен анализ. От такъв индексен анализ се получават **положителни съвместни ефекти** от едновременните намаления на двата фактора (!), както и **съвместни ефекти** от разнопосочните промени на факторите! Мисля, че е излишен всякакъв коментар на такава логика, но по-подробно за нея читателят ще намери в моята предходна статия в списанието през 2016 година. Тя е за индексния факторен анализ на **продукцията** от еднородни съвкупности на стоките според промените на тяхната средна цена \bar{p} и на общото натурално количество Q с **дискретната функция на математическия сигнум** (Христов, 2016б). Други автори предлагат по две решения за адитивния факторен анализ, но **без аналитичен или логически критерий** за вярното еднозначно решение, който произлиза от дискретната функция на математическия сигнум. Както ще се види от следващото изложение, **неверни** са решенията от икономическото образование и на измененията на **средните равнища (цени)** от промените на цените на отделните стоки Δp_i и на относителните дялове $\Delta \omega_i$ на техните



натурални количества. Това изложение се отнася също и за адитивния факторен анализ на демографските процеси, по-конкретно на разликите между броя на умрелите мъже и жени по възраст, както и за изменението на общата смъртност на населението.

Цитираните автори от икономическото образование, някои от които не са демографи, предлагат различни, но общо **четири метода**, ако анализираха същата разлика за общата смъртност $\Delta \bar{m} = \bar{m}_1 - \bar{m}_0$ от промените на коефициентите на смъртността на отделните възрасти $\Delta m_i = m_{i1} - m_{i0}$ и на относителните дялове на населението $\Delta \omega_i = \omega_{i1} - \omega_{i0}$.

Първият метод е с трите сумарни ефекта с тегла от базисната година:

$$\Delta \bar{m} = \bar{m}_1 - \bar{m}_0 = \Delta \bar{m}, m + \Delta \bar{m}, \omega + \Delta \bar{m}, m\omega = \sum(m_{i1} - m_{i0}) \omega_{i0} + \sum(\omega_{i1} - \omega_{i0}) m_{i0} + \sum(m_{i1} - m_{i0}) (\omega_{i1} - \omega_{i0})$$
 [Цонев, 1968; Казинец, 1969; Къналиев, 1975; Сугарев, 1975; Гатев, 1995; Русев и Сугарев, 2008; Стойкова-Къналиева, 2016, Аркадиев, 2017].

За проф. В. Цонев и проф. З. Сугарев този метод е най-съдържателен за адитивен факторен анализ на средни равнища, защото с него се получавали три ефекта. Структурният ефект $\sum(\omega_{i1} - \omega_{i0}) m_{i0}$ се извежда от структурния индекс $I_{str} = \frac{\sum \omega_{i1} m_{i0}}{\sum \omega_{i0} m_{i0}}$ чрез разликата $\sum \omega_{i1} m_{i0} - \sum \omega_{i0} m_{i0}$. Той се появява за първи път в нашата икономическа литература в превод от чужд източник - Ройс (Reyss, немски автор), в който вместо m_{i0} се използва y_{i0} - отраслова производителност на труда на едно заето лице от базисната година (Къналиев, 1975).

С вторите два метода се получават решения, които съдържат само двата сумарни ефекта $\Delta \bar{m}, m$ и $+\Delta \bar{m}, \omega$ без сумарния съвместен ефект $\Delta \bar{m}, m\omega$:

$$\Delta \bar{m} = \bar{m}_1 - \bar{m}_0 = \Delta \bar{m}, m + \Delta \bar{m}, \omega = \sum(m_{i1} - m_{i0}) \omega_{i0} + \sum(\omega_{i1} - \omega_{i0}) m_{i1}$$

$$\Delta \bar{m} = \bar{m}_1 - \bar{m}_0 = \Delta \bar{m}, m + \Delta \bar{m}, \omega = \sum(m_{i1} - m_{i0}) \omega_{i1} + \sum(\omega_{i1} - \omega_{i0}) m_{i0}$$
 [Перегудов, 1960; Цонев, 1968, 1970; Сугарев, 1975; Гатев, 1995; Къналиев, 2005; Стойкова-Къналиева, 2016].

Посоченият адитивен ефект от структурните промени в първото решение $\sum(\omega_{i1} - \omega_{i0}) m_{i1}$ е известен за първи път у нас също от чужд източник (Перегудов, 1960). Той също се извежда от структурен индекс $I_{str} = \frac{\sum \omega_{i1} m_{i1}}{\sum \omega_{i0} m_{i1}}$ чрез разликата $\sum \omega_{i1} m_{i1} - \sum \omega_{i0} m_{i1}$, където m_{i1} е средна икономическа характеристика по отрасли или по друг

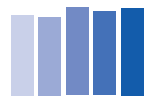
признак. По-нататък, като възприема този структурен индекс, доц. Къналиев го подразделя на три факторни структурни субиндекса (Къналиев, 1975, с. 259). От своя страна, проф. Гатев подразделя същия структурен индекс само на два факторни субиндекса (Гатев, 1995, с. 364, 372). Моето отношение към такива индекси и адитивни структурни ефекти е отрицателно. Според предложението адитивен факторен анализ в настоящата статия тези ефекти не могат да се измерват с умножения на структурните промени $\Delta\omega_i$ **само** с другия фактор p_{i0} от базисната година, нито **само** с p_{i1} от отчетната година. След като нетните и общите структурни ефекти са условни, неточни и следователно неверни, също такива са и техните факторни части, на които ги подразделят цитираните автори.

В последното преработено издание на учебника „Демографска статистика“ от професорите Б. Русев и З. Сугарев проф. Русев е изложил два метода за решения със сумарните съвместни ефекти. Първото решение е посоченото по-горе с трите сумарни ефекта и тегла от базисната година. С втория метод решението е с тегла от отчетната година:

$$\bar{m} = \bar{m}_1 - \bar{m}_0 = \Delta\bar{m}, m + \Delta\bar{m}, \omega + \Delta\bar{m}, m\omega = \sum (m_{i1} - m_{i0}) \omega_{i1} + \sum (\omega_{i1} - \omega_{i0}) m_{i1} + \sum (m_{i1} - m_{i0})(\omega_{i1} - \omega_{i0}) [\text{Русев, Б. и З. Сугарев, 2008; Русев, 2009}].$$

За тези решения важи същата критика както за предходните решения. Специално за второто решение **нима е възможно при едновременни намаления на двата фактора $\Delta m_i < 0$ и $\Delta \omega_i < 0$ да се получават положителни съвместни ефекти?** Или обобщено, **всички ефекти** в представените четири решения се измерват чрез произведенията на промените на единия фактор със стойностите на другия фактор или **само** от базисната година, или **само** от отчетната година! Следователно за **някои възрасти ефектите ще бъдат неверни**, а чрез тях и **общите (сумарни) ефекти**, защото се използват по-големите стойности на другия фактор, а не по-малките според дискретната функция на математическия сигнум!

От съвременните автори е цитиран колектив с ръководител доц. Андреана Стойкова-Къналиева, който е извършил едно голямо и задълбочено международно изследване на бруtnата добавена стойност за България и за останалите страни на Европейския съюз за периода 2000 - 2014 г. (Стойкова-Къналиева, А., А. Найденов и В. Бозев, 2016). В това всестранно изследване са анализирани както бруtnата добавена стойност като **обемна резултативна величина** от промените на нейното средно равнище на едно заето лице и на броя на заетите лица, така и на средното равнище като



претеглена средна от секторните (отраслови) средни равнища и относителните дялове на заетите по сектори (отрасли). Показано е, че адитивният факторен анализ на изменението на брутната добавена стойност като обемна резултативна величина може да се извърши по най-разгърнатата схема на анализ със седем ефекта. Те се получават от разгърнат адитивен факторен анализ на изменението на средното равнище на този показател с трите сумарни ефекта, но **с тегла само от базисния период!** Отделно е отбелязано, че специален интерес в сравнителните териториални изследвания можело да представлява решението на разликата на средните равнища и **само** с двата сумарни ефекта $\Delta \bar{m}, m = \sum (m_{i1} - m_{i0}) \omega_{i0}$ и $\Delta \bar{\omega}, \omega = \sum (\omega_{i1} - \omega_{i0}) m_{i1}$ (с. 54).

Моето мнение е, че не може да се преминава от един модел и метод за адитивен факторен анализ в друг модел и метод с едни и същи данни, без логически и аналитичен критерий за избора на даден метод. Излиза, че според проф. Сугарев, както и според доц. Стойкова-Къналиева, това преминаване било необходимо **само за да се използвали тегла от един и същ период ω_{i0} и m_{i0} (?)** (Сугарев, 1975, с. 97).

Доц. Стойкова-Къналиева е предложила за решение на адитивния факторен анализ на обемните резултативни величини с агрегираните (обобщаващи) данни за средното равнище (интензивния показател) и за общото натурално количество (екстензивния показател) следния адитивен факторен модел:

$\Delta V = V_1 - V_0 = \Delta \bar{y} \times N_1 + \Delta N \times \bar{y}_0$, където V_1 и V_0 са обеми на брутната добавена стойност, N е броят на заетите и \bar{y}_0 - средната брутна добавена стойност на едно заето лице от базисната година¹ (Стойкова-Къналиева и др., 2016, с. 49).

Този модел обаче според теорията на вероятностите и дискретната нечетна функция на математическия сигнум е верен **само** при разнопосочните факторни промени $\Delta \bar{y}_0 > 0$ и $\Delta N < 0$. За останалите три вида факторни промени моделът не е верен! Например според данните от изследването неблагоприятна за икономиката на България, както и за други страни, е била 2009 г., за която има **едновременни намаления** на брутната добавена стойност, броя на заетите и на средната брутна добавена стойност на едно заето лице спрямо предходната 2008 година (Стойкова-Къналиева и др., 2016, Приложение В, табл. В-1, с. 198, и табл. В-2, с. 200).

След това доц. Стойкова-Къналиева разширява адитивния факторен анализ на разликата ΔV също с **неверен модел и метод** за адитивен анализ на отрицателната

¹ Означенията са според цитирания източник.

разлика на двете средни брутни добавени стойности на едно заето лице $\Delta\bar{y} = \bar{y}_1 - \bar{y}_0 = -76.1$ хиляди евро:

$$\Delta V = N_0 \sum d_{i0} \Delta\bar{y}_i + \Delta N \sum d_{i0} \Delta\bar{y}_i + N_0 \sum \Delta d_i \bar{y}_{i0} + \Delta N \sum \Delta d_i \bar{y}_{i0} + N_0 \sum \Delta d_i \Delta\bar{y}_i + \Delta N \sum \Delta d_i \Delta\bar{y}_i + \Delta N \sum d_{i0} \bar{y}_{i0},$$

където d_{i0} е относителният дял на заетите в i -тия сектор (отрасъл) от базисната година;

$\Delta d_i = d_{i1} - d_{i0}$ е промяната на относителния дял на заетите в i -тия сектор (отрасъл) през отчетната спрямо базисната година;

\bar{y}_{i0} е средната брутна добавена стойност на едно заето лице в i -тия сектор (отрасъл) през базисната година;

$\Delta\bar{y}_i = \bar{y}_1 - \bar{y}_0$ е промяната на средната добавена стойност на едно заето лице в i -тия сектор (отрасъл) през отчетната спрямо базисната година (Пак там, с. 51).

Внимателният читател ще забележи, че всички промени на единия фактор в този модел се умножават също със стойностите на другия фактор **само** от базисната година. Извинявам се за откровеността, но по тази причина отхвърлям категорично всички подобни факторни анализи. За съжаление, по същата причина отхвърлям моделите и методите и на останалите цитирани автори от катедра „Статистика и иконометрия“ на УНСС, някои от които са били мои преподаватели. Проблемът обаче е много по-голям и стар, защото засяга не само икономическото образование и обществените науки, икономическата и демографската статистика, но и **приложенията на теорията на вероятностите**. Адитивният факторен анализ на средните величини е **първият масов и най-необходим** дискретен статистически факторен анализ във всички области на живота.

Заклучение

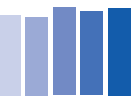
С новата методика за адитивен факторен анализ на разликата на две средни цени на стоките от еднородни съвкупности с неагрегирана информация за отделните стоки завършва поредицата от авторови методики за този анализ (Христов, 2015, 2016а). Те са три предходни методики за изменението на продукцията (дискретната зависима променлива) като обемна резултативна величина в паричен израз от промените на дискретните факторни променливи - цените и натуралните количества на стоките за две сравнявани години (базисна и отчетна). Първата методика е за изменението на продукцията само на отделната стока (Христов, 2015). Другите две методики са също за изменението на продукцията, но отделно от еднородни и от разнородни съвкупности на стоките (Христов, 2016а).

Новата методика е за изменението на средната цена (дискретната зависима променлива) от промените на дискретните факторни променливи - цените на отделните стоки от еднородните съвкупности и на относителните дялове на техните **сравними** натурални количества. Следователно адитивният факторен анализ на средните цени представлява **продължение** на адитивния анализ на обемните резултативни величини (продукцията) от еднородните съвкупности на стоките с тяхната агрегирана информация. Оттук произлиза и **универсалността** на новата методика за адитивен факторен анализ в приложните статистики от всички области на живота, които работят с еднородни статистически съвкупности и неагрегирана информация за тях.

Новото в предлаганата методика са **осемте** възможни еднозначни или единствено верни и точни решения на разликата на две средни равнища от промените на показателите за интензивност на всяко явление или процес и на относителните дялове на екстензивните показатели, от които произлиза явлението или процесът. Второто **ново** в методиката е, че в **общия случай** от четирите вида промени на двата фактора - еднопосочните едновременни увеличения или намаления на факторите, и от разнопосочните техни промени (единият фактор се увеличава, а другият намалява) всяко еднозначно решение съдържа **три ефекта** - $\Delta\bar{p}, p$, $\Delta\bar{p}, \omega$ и $\Delta\bar{p}, p\omega$. Те са едновременни увеличения и/или намаления на средното базисно равнище (цената \bar{p}_0). Двата от тях $\Delta\bar{p}, p$ и $\Delta\bar{p}, \omega$ са **нетни** само от промяната на всеки фактор, докато третият ефект $\Delta\bar{p}, p\omega$ е **съвместен**, макар и минимален, само от еднопосочните промени на двата фактора. Единият от нетните ефекти $\Delta\bar{p}, p$ е **интензивният** (ценовият) само от промените на интензивните факторни показатели (цените на отделните стоки), а

другият нетен ефект е **структурният** $\Delta \bar{r}, \omega$ само от структурните промени или разликите на относителните дялове на екстензивните показатели (натуралните количества на стоките). Третото **ново** в методиката е, че **всички ефекти** от адитивния факторен анализ на средните равнища както ефектите от адитивния анализ на обемните резултативни величини могат да се изведат не само с индуктивната обобщаваща логика или с нейния аналитичен аналог - дискретната функция на математическия сигнум, но и с **теорията на вероятностите**. Приложенията на тази теория за статистически анализи са всички дискретни статистически адитивни и индексни факторни анализи във всички икономически и неикономически статистики. Тук е необяснима ролята и участието на математиката не само в икономическата статистика, но и във всички други неикономически приложни статистики като демографската, трудовата, социалната и застрахователната статистика. Всички статистики прилагат адитивния факторен анализ като **първия и елементарен, но най-необходим на практиката и научно-приложните изследвания** дискретен статистически факторен анализ. Ако статистическите задачи в икономиката са **принципно различни** за еднородните и разнородните съвкупности и са първо **икономически задачи**, а след това математически, всички други задачи са за еднородни съвкупности и са следователно **математически**. Техният адитивен факторен анализ е всъщност едно от най-важните приложения на теорията на вероятностите. Абсолютно е необяснимо обаче как досега този анализ не е изведен с тази теория от математици, които я владеят и се занимават или преподават демографска статистика, човешки ресурси или животозастраховане. Към тази критика ще добавя и **общата** към математическото образование в икономическите колежи и висши училища, че не се преподава дискретната нечетна функция на математическия сигнум за определяне на верните и точни ефекти от всички адитивни факторни анализи на обемни резултативни величини и на средни равнища.

С новата методика за адитивен факторен анализ се отхвърлят всички други традиционни методики за този анализ в икономическото образование и обществените науки у нас и в чужбина. Причината е, че с традиционните методики **всеки ефект** се получава като произведение на промяната на всеки фактор със стойността на другия фактор **само** от базисната година или **само** от отчетната година, а не с неговата **по-малка стойност** според дискретната функция на математическия сигнум. Получените по този начин ефекти за продукцията, както и за средните цени с традиционните методики, ще бъдат **брутни** за някои стоки. По-конкретно, при анализа на средните цени те ще съдържат освен двата верни нетни ефекта още и **два фиктивни (реално**



допълнително увеличение на средната базисна цена \bar{p}_0 през отчетната година, а другият фиктивен ефект е със същия размер, но с отрицателен знак. Той означава **едновременно допълнително намаление** на същата средна базисна цена \bar{p}_0 през същата отчетна година, или двата фиктивни ефекта са **логически абсурдни като допълнителни ефекти!** Дори само от формална гледна точка те взаимно се анулират и не са ефекти.

Една следваща статия ще бъде за много по-трудния индексен факторен анализ на средните равнища, който се извежда с верните и точни ефекти от адитивния факторен анализ в настоящата статия с дискретната функция на математическия сигнум.

Приложение 1

Цени и натурални количества на стоките и техните влияния върху изменението на обема на продукцията

Стоки	Базисна година			Отчетна година			Ефекти от адитивния анализ			
	P_{i0}	Q_{i0}	$P_{i0}Q_{i0}$	P_{i1}	Q_{i1}	$P_{i1}Q_{i1}$	$\Delta p_i q_{im}$	$\Delta q_i p_{im}$	$h_i \Delta p_i \Delta q_i$	$P_{i1}Q_{i1} - P_{i0}Q_{i0}$
	хил. лв.	бр.	хил. лв.	хил. лв.	бр.	хил. лв.	хил. лв.	хил. лв.	хил. лв.	хил. лв.
	1	2	3=1x2	4	5	6=4x5	7	8	9	10=7+8+9
А	40	40	1600	80	50	4000	+1600	+400	+400	+2400
Б	60	50	3000	50	30	1500	-300	-1000	-200	-1500
В	50	50	2500	90	30	2700	+1200	-1000	-	+200
Г	60	60	3600	50	70	3500	-600	+500	-	-100
Общо	53.5	200	10700	65.0	180	11700	+1900	-1100	+200	+1000

Приложение 2

Цени и натурални количества на стоките и техните влияния върху изменението на обема на продукцията

Стоки	Базисна година			Отчетна година			Ефекти от адитивния анализ			
	P_{i0}	Q_{i0}	$P_{i0}Q_{i0}$	P_{i1}	Q_{i1}	$P_{i1}Q_{i1}$	$\Delta p_i q_{im}$	$\Delta q_i p_{im}$	$h_i \Delta p_i \Delta q_i$	$P_{i1}Q_{i1} - P_{i0}Q_{i0}$
	хил. лв.	бр.	хил. лв.	хил. лв.	бр.	хил. лв.	хил. лв.	хил. лв.	хил. лв.	хил. лв.
	1	2	3=1x2	4	5	6=4x5	7	8	9	10=7+8+9
1	80	50	4000	40	40	1600	-1600	-400	-400	-2400
2	50	30	1500	60	50	3000	+300	+1000	+200	+1500
3	90	30	2700	50	50	2500	-1200	+1000	-	-200
4	50	70	3500	60	60	3600	+600	-500	-	+100
Общо	65.0	180	11700	53.5	200	10700	-1900	+1100	-200	-1000

Приложение 3

Цени и натурални количества на стоките и техните влияния върху изменението на обема на продукцията

Стоки	Базисна година			Отчетна година			Ефекти от адитивния анализ			
	P_{i0}	Q_{i0}	$P_{i0}Q_{i0}$	P_{i1}	Q_{i1}	$P_{i1}Q_{i1}$	$\Delta p_i q_{im}$	$\Delta q_i p_{im}$	$h_i \Delta p_i \Delta q_i$	$P_{i1}Q_{i1} - P_{i0}Q_{i0}$
	хил. лв.	бр.	хил. лв.	хил. лв.	бр.	хил. лв.	хил. лв.	хил. лв.	хил. лв.	хил. лв.
	1	2	3=1x2	4	5	6=4x5	7	8	9	10=7+8+9
1	90	90	8100	100	100	10000	+900	+900	+100	+1900
2	80	80	6400	90	85	7650	+800	+400	+50	+1250
3	80	50	4000	60	35	2100	-700	-900	-300	-1900
4	40	20	800	30	10	300	-100	-300	-100	-500
5	20	30	600	60	20	1200	+800	-200	-	+600
6	80	20	1600	50	30	1500	-600	+500	-	-100
Общо	74.138	290	21500	81.250	280	22750	+1100	+400	-250	+1250

Приложение 4

Цени и натурални количества на стоките и техните влияния върху
изменението на обема на продукцията

Стоки	Базисна година			Отчетна година			Ефекти от адитивния анализ			
	P_{i0} хил. лв.	Q_{i0} бр.	$P_{i0}Q_{i0}$ хил. лв.	P_{i1} хил. лв.	Q_{i1} бр.	$P_{i1}Q_{i1}$ хил. лв.	$\Delta p_i q_{im}$ хил. лв.	$\Delta q_i p_{im}$ хил. лв.	$h_i \Delta p_i \Delta q_i$ хил. лв.	$P_{i1}Q_{i1} - P_{i0}Q_{i0}$ хил. лв.
	1	2	3=1x2	4	5	6=4x5	7	8	9	10=7+8+9
1	100	100	10000	90	90	8100	-900	-900	-100	-1900
2	90	85	7650	80	80	6400	-800	-400	-50	-1250
3	60	35	2100	80	50	4000	+700	+900	+300	+1900
4	30	10	300	40	20	800	+100	+300	+100	+500
5	60	20	1200	20	30	600	-800	+200	-	-600
6	50	30	1500	80	20	1600	+600	-500	-	+100
Общо	81.250	280	22750	74.138	290	21500	-1100	-400	+250	-1250

Приложение 5

Цени и натурални количества на стоките и техните влияния върху
изменението на обема на продукцията

Стоки	Базисна година			Отчетна година			Ефекти от адитивния анализ			
	P_{i0} хил. лв.	Q_{i0} бр.	$P_{i0}Q_{i0}$ хил. лв.	P_{i1} хил. лв.	Q_{i1} бр.	$P_{i1}Q_{i1}$ хил. лв.	$\Delta p_i q_{im}$ хил. лв.	$\Delta q_i p_{im}$ хил. лв.	$h_i \Delta p_i \Delta q_i$ хил. лв.	$P_{i1}Q_{i1} - P_{i0}Q_{i0}$ хил. лв.
	1	2	3=1x2	4	5	6=4x5	7	8	9	10=7+8+9
1	40	40	1600	80	60	4800	+1600	+800	+800	+3200
2	60	50	3000	40	40	1600	-800	-400	-200	-1400
3	40	50	2000	60	40	2400	+800	-400	-	+400
4	60	40	2400	50	60	3000	-400	+1000	-	+600
Общо	50	180	9000	59	200	11800	+1200	+1000	+600	+2800

Приложение 6

Цени и натурални количества на стоките и техните влияния върху
изменението на обема на продукцията

Стоки	Базисна година			Отчетна година			Ефекти от адитивния анализ			
	P_{i0} хил. лв.	Q_{i0} бр.	$P_{i0}Q_{i0}$ хил. лв.	P_{i1} хил. лв.	Q_{i1} бр.	$P_{i1}Q_{i1}$ хил. лв.	$\Delta p_i q_{im}$ хил. лв.	$\Delta q_i p_{im}$ хил. лв.	$h_i \Delta p_i \Delta q_i$ хил. лв.	$P_{i1}Q_{i1} - P_{i0}Q_{i0}$ хил. лв.
	1	2	3=1x2	4	5	6=4x5	7	8	9	10=7+8+9
1	80	60	4800	40	40	1600	-1600	-800	-800	-3200
2	40	40	1600	60	50	3000	+800	+400	+200	+1400
3	60	40	2400	40	50	2000	-800	+400	-	-400
4	50	60	3000	60	40	2400	+400	-1000	-	-600
Общо	59	200	11800	50	180	9000	-1200	-1000	-600	-2800

ЦИТИРАНА ЛИТЕРАТУРА:

Аркадиев, Д. (2017). Някои структурни промени в родилния (фертилния) контингент в България и влиянието им върху раждаемостта, Статистика, кн. 3, www.nsi.bg.

Гатев, К. (1995). Въведение в статистиката, Лиа, С.

Казинец, Л. (1969). Измерение структурных сдвигов в экономике, Экономика, Москва.

Къналиев, Т. (1975). Структурните промени в народното стопанство и повишаване обществената производителност на труда - измерване, анализ, перспективи, С., Трудове на ВИИ „Карл Маркс“, книга 1.

Къналиев, Т. (2005). Съвкупностният (статистическият) подход за изследване и някои дискуссионни въпроси на индексната теория, Статистика, кн. 1, С.

Перегудов, В. (1960). Теоретические вопросы индексного анализа, Статистика, Москва.

Русев, Б., З. Сугарев (2008). Демографска статистика, Университетско издателство „Стопанство“, С.

Русев, Б. (2009). Един подход за адитивен индексен анализ, Икономическа мисъл, кн. 5, С.

Стойкова-Къналиева, А., А. Найденов и В. Бозев (2016). Статистическо сравнително изследване на структурите, структурните различия и структурната динамика на основни макроикономически показатели (брутна добавена стойност) на страните от Европейския съюз през периода 2000 - 2014 г., Издателски комплекс - УНСС, София, 2016.

Сугарев, З. (1975). Демографска статистика, Наука и изкуство, С.

Христов, Е. (1978). Прирастът на продукцията според промените на вложеното количество труд и производителността на труда, Статистика, кн. 5, С.

Христов, Е. (1983). Адитивни методи за оценки на влияния и приноси на народностопанските отрасли върху обществената производителност на труда, Статистика, кн. 5, С.

Христов, Е. (1986). Ефекти от промените на смъртността и структурата на населението по възраст, Население, кн. 2, С.

Христов, Е. (1987). Оценяване на структурни и неструктурни ефекти в икономиката, Икономика, кн. 8, С.

Христов, Е. (2004а). Факторен анализ на прирасти на средни равнища с реални нетни и брутни ефекти, Икономическа мисъл, кн. 6, С.

Христов, Е. (2004б). Анализ на изменението на производителността на труда с реални структурни и неструктурни ефекти, Статистика, кн. 6, С.

Христов, Е. (2008). Едно достатъчно условие за еднозначни решения на факторни промени на средни равнища, Статистика, кн. 4, С.

Христов, Е. (2013). Факторни модели за общото, прякото и косвеното влияние на повъзрастовата смъртност върху изменението на средната продължителност на живота, Статистика, кн. 3 - 4, С.

Христов, Е. (2015). Елементарният функционален адитивен и индексен факторен анализ и неговите еднозначни решения с дискретната нечетна функция на математическия сигнум, Статистика, кн.1, www.nsi.bg.

Христов, Е. (2016a). Адитивен факторен анализ на обема на продукцията на еднородни и разнородни съвкупности на стоки с дискретната нечетна функция на математическия сигнум, Статистика, кн. 1, www.nsi.bg.

Христов, Е. (2016б). Индексен факторен анализ на обема на продукцията от еднородни съвкупности на стоки с дискретната нечетна функция на математическия сигнум, Статистика, кн. 2, www.nsi.bg.

Христов, Е. (2017). Индексен факторен анализ на обема на продукцията от еднородни и разнородни съвкупности на стоки от промените на техните цени и натурални количества с дискретната нечетна функция на математическия сигнум, Статистика, кн. 3, www.nsi.bg.

Цветков, С. (2015). За някои от проблемите при статистическото оценяване на инфлацията, Национална научна конференция, посветена на Международната година на статистиката 2013, Издателски комплекс - УНСС, София.

Цонев, В. (1968). Конкретен или абстрактен подход при статистическия анализ на прираста на обема на продукцията, Трудове на ВИИ „Карл Маркс“, том I, С.

Цонев, В. (1970). За по-конкретно дефиниране на факторите при анализ на прираст на обеми, Статистика, кн. 4, С.

Цонев, В. (1977). За двете нива на статистическия диагностичен анализ, Статистика, кн. 5, С.

Югенбург, С. (1955). О разложении абсолютных приростом по факторам, Ученые записки по статистике, АН СССР, том I, М.

Янкова, Н. (2007). Статистическо изследване на структурни изменения, Академично издателство „Проф. Марин Дринов“, БАН, София.

Hristov, E. (2008). A Complex Model for Measurement and Factor Analysis of Population Ageing, Economic Thought, year XXIII, Bulgarian Academy of Sciences Institute of Economics.

The Oxford English Dictionary 1993, Oxford University Press.

The Oxford Paperback Dictionary 1994, Oxford University Press.

**АДИТИВЕН ФАКТОРЕН АНАЛИЗ НА СРЕДНИТЕ ЦЕНИ ОТ
ЕДНОРОДНИ СЪВКУПНОСТИ НА СТОКИ С
ДИСКРЕТНАТА НЕЧЕТНА ФУНКЦИЯ НА
МАТЕМАТИЧЕСКИЯ СИГНУМ**

*Емил Христов**

РЕЗЮМЕ Представена е **нова методика** за адитивен факторен анализ на изменението на средните цени на стоки от еднородни съвкупности. Само такива съвкупности имат икономически обосновани средни цени и общи натурални количества на всички стоки. Средната цена се представя с мултипликативен модел с два фактора - цените на отделните стоки и относителните дялове на техните натурални количества. Адитивният факторен анализ е на промяната на средната цена от промените на групирани данни на стоките по цени и относителните дялове на техните натурални количества. Предимството на новата методика е, че с **дискретната нечетна функция на математическия сигнум** се получават три верни и точни ефекта - **ценови** (само от промените на цените), **структурен** (само от промените на относителните дялове) и **съвместен ефект** (само от еднопосочните съвместни промени на цените и относителните дялове). От разнопосочните факторни промени няма съвместни ефекти.

Новата методика за адитивен факторен анализ е **универсална** и може да се прилага не само в икономическата статистика, но и във всички други приложни статистики, които работят с еднородни съвкупности.

* Професор, д.ик.н.; e-mail: emil_hrystov_37@hotmail.com.

АДДИТИВНЫЙ ФАКТОРНЫЙ АНАЛИЗ СРЕДНИХ ЦЕН НА ТОВАРЫ В ОДНОРОДНЫХ СОВОКУПНОСТЯХ ПРИ ИСПОЛЬЗОВАНИИ ДИСКРЕТНОЙ НЕЧЕТНОЙ ФУНКЦИИ МАТЕМАТИЧЕСКОЙ СИГМЫ

*Емил Христов**

РЕЗЮМЕ Представлена новая методика аддитивного факторного анализа изменения средних цен на товары в однородных совокупностях. Такие совокупности имеют экономически обоснованные средние цены и общие количества товаров в натуральном выражении. Средняя цена представлена мультипликативной моделью с двумя факторами: цены на отдельные товары и относительные доли их количества в натуральном выражении. Аддитивный факторный анализ представляет изменение средней цены в зависимости от изменения сгруппированных данных о товарах по ценам и относительным долям их натуральных количеств. Преимущество новой методики состоит в том, что применение дискретной нечетной функции математической сигмы позволяет получить три точных эффекта - ценовой (только в результате изменения цен), структурный (только в результате изменения относительных долей) и совместный эффект (только в результате однонаправленных совместных изменений цен и относительных долей). При разнонаправленных факторных изменениях совместные эффекты отсутствуют. Новая методика аддитивного факторного анализа является универсальной и может применяться не только в экономической статистике, но и в других областях прикладной статистики, использующих однородные совокупности.

* Д-р экон. наук, профессор.; e-mail: emil_hristov_37@hotmail.com.

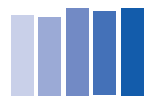
**ADDITIVE FACTORIAL ANALYSIS OF THE AVERAGE PRICES OF
HOMOGENEOUS SETS OF GOODS WITH
THE DISCRETE NON - ODD FUNCTION OF THE
MATHEMATICAL SIGNUM**

*Emil Hristov**

SUMMARY A **new methodology** for additive factorial analysis of the change in average prices of goods from homogeneous clusters is presented. Only such aggregates have economically justified average prices and total natural quantities of all commodities. The average price is represented by a multiplier model with two factors - prices of the individual goods and relative shares of their natural quantities. The additive factorial analysis is the change in the average price of the changes in grouped data of goods by prices and relative shares of their natural quantities. The advantage of the new methodology is that **with the discreet odd function of the mathematical signum** – three correct and precise effects are obtained - **price** (only from price changes), **structural** (only from changes in relative shares) and **joint effect** (only by unidirectional joint changes of prices and relative shares). There are no joint effects from the unidirectional joint changes.

The new methodology for additive factorial analysis is **universal** and can be applied not only in economic statistics but also in all other applied statistics that work with homogeneous clusters.

* Professor, PhD; e-mail: emil_hristov_37@hotmail.com.



ПРИЛОЖНИ АСПЕКТИ НА „ГОЛЕМИТЕ ДАННИ“ В ОФИЦИАЛНАТА СТАТИСТИКА

*Галя Статева**



Въведение

Процесите на глобализация и технологизация във всички сфери на обществения живот оказват огромен натиск по отношение на управлението в национален и международен аспект. За успешното управление е необходима „добра“ информация, което означава, тя да притежава следните характеристики: синтезирана, кратка, точна, акцентирана, подсказваща най-доброто решение на проблемите и очертаваща (прогнозираща) хоризонтите за развитие на процесите и явленията. В този смисъл Big Data е друга, различна от тази, която познаваме, екосистема. Анализът на основата на „големите данни“ може да бъде отвъд източниците, измерването и напрежението, необходими при създаването на информация, а също така отвъд политиката. Успоредно с това през 21-ви век тази екосистема е исторически феномен на човешкото развитие.

В последните няколко години редица национални статистически офиси осъществяват европейски и международни проекти за „големите данни“. Опитът от успешното използване на „големите данни“ може да бъде изучаван и споделян с други държави с цел извличане на ценни познания и прилагане на добри практики по отношение на Big Data. Освен това националните статистически организации са окуражавани от Евростат да включат официално въпросите и за

* Д-р, държавен експерт в отдел „Обща методология и анализ на статистическите изследвания“, дирекция „Методологично-учебен център“, НСИ; e-mail: GStateva@nsi.bg.

„големите данни“ в техните годишни програми и стратегически документи чрез осъществяване на експериментално-изследователски и пилотни проекти в избрани области и чрез разпределяне на подходящите ресурси за тези цели. В процеса на приложение на Big Data в официалната статистическа практика Евростат играе ключова роля, като се има предвид факта, че той е източник на финансови, идейни и технически ресурси, включително и организатор на редица специализирани обучения, семинари и научни конференции по тази тематика.

Националните статистически институти създават разнообразни масиви от статистически данни за използване на информационните и комуникационните технологии, които се използват за наблюдение на напредъка на страните към информационното общество. Традиционно тези данни се събират чрез два различни въпросника - един за домакинствата/физическите лица и един за предприятията. Бързото развитие на съвременните ИКТ поставя необходимостта от разработване на показатели, които да са повече релевантни и навременни от тези, изчислявани на базата на традиционните изследвания.

Дигиталните отпечатъци/следи, резултат от нашето ежедневие, могат ефективно да бъдат използвани за измерване на голямо разнообразие от явления. В проведено проучване през 2013 г. Европейската комисия изучава възможността за използване на интернет като допълнителен източник на данни или дори като заместител на традиционните статистически източници¹. Като резултат от изследването се стига до заключението, че относно текущите променливи за физически лица/домакинства нито достъпът до ИКТ, нито използването на компютри са възможни за измерване чрез интернет. От друга страна, прогнозируемото използване на интернет мрежата от предприятията е идеално за измерване чрез интернет. Налага се мнението, че събраните данни по електронен път биха предоставили достатъчно материал за изчисляване на допълнителни индикатори за ИКТ, както и голям брой потенциални показатели, които предлагат допълнителни възможности за изучаване на съществуващи явления и необхванати в досегашното традиционно статистическо изследване за ИКТ в предприятията.

През 2013 г. Италианският статистически институт (ISTAT) започва да тества техники за извличане на информация от интернет (web-scraping) и извличане на съдържание от текст (text mining) за около 5 600 уебсайта от общо 8 600 адреса на уебсайтове, посочени от предприятията при попълване на въпросника на традиционното изследване, като събраното съдържание се индексира чрез технологични инструменти и алгоритми и се оценяват резултатите от

¹ Internet as data source. Feasibility Study on Statistical Methods on Internet as a Source of Data Gathering. (2012). European Commission, SMART 2010/30. file:///D:/Fullfinalreportofthestudy.pdf.



статистическото изследване. По този начин ISTAT прави практически опит за промяна на традиционното изследване „Използване на ИКТ в предприятията“ (Information and Communication Technologies in Enterprises)² чрез осигуряване на допълнителни източници на данни. Заключениета и съображенията от опита на ISTAT относно използването на интернет като източник на данни за официалната статистика са проучени от българските специалисти и са основа, на която те проектират настоящото изследване.

В началото на 2016 г. Националният статистически институт на Р България (НСИ) се включи като страна партньор в ESSnet проект „Рамково споразумение за сътрудничество Big Data План за действие“, който е разделен на две отделни грантови споразумения (SGA-I и SGA-II), изпълняващи се последователно за периода 2016 - 2018 година. В рамките на това споразумение се постигат целите, заложи в т.нар. BDAR (Big Data Action Plan and Roadmap v. 1.0), който е част от портфолиото на ECC Vision 2020. Основната цел на проекта е да подготви Европейската статистическа система за интегриране на източници на „големи данни“ в процеса на производство на официална статистика. Специфичните цели на работните пакети са свързани с анализиране на получените резултати от източниците на Big Data; разработване на методология за използването на Big Data в статистическата практика и измерване на качеството на събраните данни; идентифициране, дефиниране и внедряване на ИТ инфраструктура за обработка и съхранение на Big Data; правни въпроси, свързани с достъпа и използването на източниците на Big Data в рамките на ECC; обмен на информация между официалната статистическата система и научната общност.

Статистиката за използването на ИКТ е естествен и логичен „кандидат“ за пилотен проект и реинженеринг на базата на интернет и подобни източници. Поради тази причина **през 2016 г. екип³ от НСИ** провежда емпирично изследване на тема „Извличане на информация от интернет за характеристики на предприятията (web-scraping)“ в рамките на европейския проект.

Основната цел на проведеното емпирично изследване е аналогична на проведеното от ISTAT изследване и е насочена към проучване на възможностите за прилагането на техниките „web-scraping“ и „text mining“ и други подобни, както и да се оцени ефектът от използването им в процеса на събиране на данни и подобряване на качеството на информацията за предприятията от бизнес регистъра на НСИ чрез достъп до техните уебсайтове. Специфичните цели на изследването са: да се демонстрира дали статистическите бизнес регистри могат да бъдат подобрени чрез използване на web-scraping и прилагане на моделно-ориентирани подходи, за да

² Information and communication technologies in enterprises. <http://www.istat.it/en/archive/77760>.

³ Авторът на статията е ръководител на екипа на НСИ и участва активно като основен статистически експерт при организацията, провеждането и анализа на резултатите от изследването.

се предскажат стойностите на някои ключови променливи за всяко предприятие; да се верифицира възможността за производство на статистически резултати от масив с големи данни с по-надеждна прогнозна стойност и съчетаването им с данни от традиционно статистическо изследване или административни данни. За сравнение и валидиране се използват данните, получени чрез статистическото изследване „Използване на ИКТ от предприятията“, което се провежда регулярно и съгласно европейските стандарти и националното законодателство от НСИ. Наблюдението е годишно, извадково, като в обхвата му на случаен принцип се включват около 4 900 предприятия. Генералната съвкупност обхваща всички предприятия от нефинансовия сектор с 10 и повече заети лица, включени в статистическия бизнес регистър на НСИ. Именно тази съвкупност е базова за провеждане на изследването за извличане на „големи данни“ от интернет чрез прилагане на техниките „web-scraping“. В списъка са включени 26 836 предприятия, които са проучени за наличие на сайтове. Първоначално в записите на бизнес регистъра са открити 2 006 броя фирмени URL адреси и 20 649 броя имейл адреси.

Работният процес на изследването е организиран в *четири основни фази*, а именно: фаза 1: Спецификация на т.нар. „сценарии“ („use-cases“), за да се дефинира обхватът на изследването; фаза 2: Разработване на един или повече пилотни проекта за всеки „use-case“, които да бъдат внедрени експериментално в НСИ; фаза 3: Практическа реализация на пилотните проекти от НСИ и фаза 4: Идентифициране на основните методологически и технологични въпроси като резултат от реализацията на пилотните проекти.

За постигане на специфичните цели на емпиричното изследване са осъществени няколко основни задачи, както следва:

Задача 1: Достъп до данни. В рамките на тази задача са извършени дейности по идентифициране на набор от методи за търсене на URL адреси за предприятия, за които те не са налични и проучване на правните аспекти относно достъпа до данни на уебсайтовете на предприятията. Проучването на правните и етичните аспекти за достъпа и съхранението на „големи данни“ беше особено важна предпоставка и условие за стартиране на действителната работа по проекта. Законодателната рамка за използване на данни от други източници за производство на официална статистика в България се състои от Закона за статистиката, Закона за защита на личните данни, Закона за авторското право и сродните му права, Закон за електронната търговия, Общ регламент за защита на данните (приложим в България от 25 май 2018 г.) и не представлява пречка за извършване на дейности по „извличане“ на информация от Интранет мрежата. По отношение на надлежно осведомяване на предприятията, които са обект на изучаване в контекста на извличане на данни от корпоративните им уебсайтове, НСИ предприе



специална кампания на сайта си, за да информира всички заинтересовани предприятия и други потребители, че ще бъде извършен масивен „web-scraping“ с експериментална цел за нуждите на официалната статистика.

Задача 2: Боравене с данните. В обхвата на тази задача се включва подробното дефиниране на „сценариите“, провеждане на консултации с потребители и други заинтересовани лица за анализиране на нуждите от статистически данни и координация с дейностите по проект ESS.VIP „Европейска система за оперативно съвместими бизнес регистри“. Проучена е научната литературата относно техниките и наличните свободни софтуерни продукти за масивен „web-scraping“ (JSoup, HTTrack и т.н.) и изучаване на проблемите, свързани с достъпността на уебсайтовете (блокиращ механизъм). Предвижда се внедряване на една или повече техники за „web-scraping“ и провеждане на същинското извличане на съдържание на уебсайтовете на предприятията и последващото им съхранение на събраното съдържание в база данни. За дефиниране на обхвата на изследването са избрани следните **use-cases**:

- **Use-case 1.** Генериране на списък с фирмени URL адреси на предприятията за бизнес регистъра (**URLs retrieval**).
- **Use-case 2.** Електронна търговия в предприятията (E-commerce) - прогнозиране дали дадено предприятие предоставя възможности за електронна търговия на фирмения си уебсайт, или не.
- **Use-case 3.** Присъствие на предприятията в социалните медии (Social media presence) - търсене и събиране на информация от фирмения уебсайт дали дадено предприятие съществува в различни социални медии.
- **Use-case 4.** Апробиране на софтуер в НСИ (разработен от италианския статистически офис ISTAT) за генериране на списък с фирмени URL адреси на предприятията (**URLs retrieval**) и сравнение на получените резултати от българския и италианския софтуер.

Задача 3: Тестване на методи и техники. За да се приложат на практика техниките „web-scraping“ е необходимо разработването на софтуерни инструменти, които се прилагат в съответствие с дефинираните „сценарии“. В рамките на тази задача екипът е разработил авторски софтуер, чрез който практически е извършено изследването. На следващо място е осигуряването на информация с цел обогатяване на статистическия бизнес регистър и изграждане на доказателства за избраните „сценарии“, за да се прогнозираат характеристики на предприятията чрез прилагане на „text mining“ и „data mining“ към събраните URL на предприятията.

Задача 4: Финализиране на методи и техники. Тази задача включва: избор на извадка от уебсайтове и ръчно определяне на свързани характеристики на предприятията и/или използване на

резултатите от изследването „Използване на ИКТ от предприятията“, за да се валидират някои характеристики на предприятията; прилагане на техники за извличане на текст и данни, за да се прогнозираат характеристики на предприятията; оценяване на „прогнозираните“ характеристики на предприятията чрез показатели за качество (например точност, чувствителност, специфичност) и сравняване и евентуално интегриране на получената информация с информацията от националните бизнес регистри.

За нуждите на изследването беше необходимо да се направи *концептуално сравнение* между етапите на класическия производствен процес за провеждане на едно традиционно изследване и основните фази на бизнес процеса за получаване на информация от източници на Big Data. За тази цел беше използвана националната версия на **Общия модел на статистическия производствен процес (ОМСПП)** в съответствие с GSBPM (Generic Statistical Business Process Model ver. 5.0)⁴, който е адаптиран към потребностите на българската статистическа система и е наличен на сайта на НСИ (http://www.nsi.bg/sites/default/files/files/metadata/NSI_GSBPM_2016.pdf). За разлика от класическото статистическо изследване процесът на получаване на информация от източници на Big Data е съществено различен, тъй като източниците на данни, достъпът до тях и средствата за тяхното събиране са напълно различни. В този смисъл може да се твърди, че GSBPM моделът е неприложим в настоящия си вид за описание на процеса на работа с Big Data. Въпреки това някои от гореописаните етапи на работа за „извличане“ на информация от интернет **могат да се съотнесат** (с определена условност) към фазите и подфазите на ОМСПП. *Първоначалното дефиниране на „сценарии“* и тяхното съдържание е съотносимо с фаза 2 от ОМСПП „Проектиране“; *задачите, свързани с достъпа до данни* от интернет източниците, съответстват на фаза 4 от ОМСПП „Събиране на данни“, която включва събирането на първичните статистически данни и зареждането им в подходяща ИТ среда за по-нататъшна обработка в контекста на традиционното статистическо изследване; *боравенето с данните от източници на Big Data* може да се свърже частично с фаза 4 и изцяло с фаза 5 от ОМСПП „Обработка на данни“, където с различни ИТ средства и техники се обработват и редактират вече събраните първични данни, като се съхраняват в подходящи бази данни; *тестването на методи и техники по отношение на вече събраните и обработени Big Data* отговаря частично на някои подфази на фаза 5 и изцяло на фаза 6 „Анализ“ от ОМСПП.

След първия опит в събирането на данни от интернет е необходимо уточнението, че съществува разграничение между автоматизираното събиране на данни (от интернет сайтове с интернет работи, „паяци“ и други подобни инструменти) без човешка намеса и асистирано

⁴ Generic Statistical Business Process Model, v5.0, (2013). <https://statswiki.unece.org/display/GSBPM/GSBPM+v5.0>.



събиране на данни (подпомагано от оператори). За втората категория е важно да се подпомага операторът, който събира данните, с цел проверка за настъпили промени в данните, достъпни на интернет сайтовете. Трябва да се отбележи обаче, че и двете категории за събиране на данни от интернет са значими и с положителни възможности за официалната статистика, включително за изучаване на социално-икономическите явления по съвсем нов начин. Автоматизираното събиране на данни може да доведе до по-детайлни данни в сравнение с данните от статистическото изследване, които могат да се използват за потвърждаване на работата на статистическите експерти, за подобряване на ефективността или за намаляване на тежестта на респондентите. От своя страна, асистираното събиране на данни би било полезно за събиране на цени на стоки и услуги от много на брой интернет сайтове по значително ефективен начин.

Идентификацията на уебсайта е друг важен въпрос, когато се използват уебсайтове за събиране на данни. Необходимо е да се оцени надеждността на сайтовете съобразно обекта на изследване, колко лесно четими са данните, кои променливи са налични и колко сравними са те сред данните от различни сайтове. Освен това трябва да се знае как нараства обемът на данните и каква е тяхната променливост - все типични характеристики за големите данни, които се събират от интернет пространството. За разлика от традиционните източници на данни (статистически въпросници, административни източници), при които характеристиките на данните са известни на организацията-доставчик или са контролирани от статистическите офиси, данните от интернет източниците са изцяло извън техния контрол.

Важен въпрос е управлението на грешките, тъй като както в традиционното изследване, така и при работа с Big Data могат да възникнат грешки на различни етапи. Някои видове грешки могат да бъдат специфични за източника; други биха могли да се прилагат за всички източници (т.е. конструктивна валидност, грешка на обхвата, грешка при измерването, грешка, дължаща се на импутация, грешка при липса на отговор от респондентите и т.н.). Извадковата грешка би могла да се приложи за специфичните случаи, където се използват извадкови методи и техники. При работа с „големи данни“ етапите за обработката им трябва да включват предварителна функция за получаване, при която данните първо се верифицират и предварително се третират, преди да последва по-задълбочен анализ, при който грешните данни, екстремалните и липсващите стойности се маркират за последваща обработка. Всички видове източници на „големи данни“ могат потенциално да са обременени от частична липса на отговор (липсващи стойности за специфични променливи). Така че в процеса за „почистване“ на „големите данни“, е необходимо да се вземе предвид, че знанието за данните и свързаните с тях метаданни са ключов фактор за разработването на ефективни методи за обработка. Предвид огромния обем на данните екстремалните стойности може да нямат влияние в сравнение с традиционната статистическа

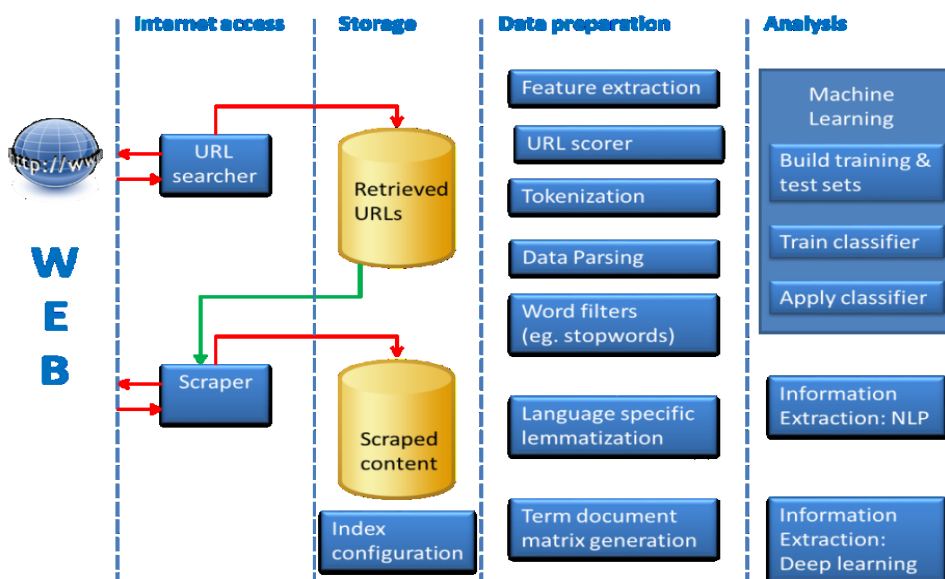
обработка. Докато методите за импутиране на данни са добре известни за традиционната статистика, все още има само няколко опита за импутиране на текстови стрингове или друг вид неструктурирани данни.

И не на последно място, огромното увеличаване на наличността на неструктурирани текстови данни изисква официалните статистически институции да увеличат финансовите си инвестиции в инструменти, способни да анализират текстови данни от интернет източници. Тези инструменти, като се започне от извличането на данни от мрежата до текстовия анализ и извличане на съдържание от текста, трябва да станат част от стандартния набор от инструменти на статистиците и анализаторите на данни.

I. Технологична среда за приложението на „web-scraping“

За изпълнение на дефинираните задачи на настоящото емперично изследване и прилагане на техниките на „web-scraping“ е разработена обща референтна логическа архитектура⁵, съставена от четири блока, съответстващи на четирите основни етапа на работата по извличането на данни от уебсайтовете на предприятия, а именно: „Интернет достъп“ (Internet access); „Съхранение“ (Storage); „Подготовка на данни“ (Data preparation) и „Анализ“ (Analysis). За всеки етап се описват логическите функционалности, които трябва да бъдат изпълнени от специфичните софтуерни продукти, разработени за нуждите на настоящото емперично изследване (фиг. 1).

Фиг. 1. Референтна логическа архитектура на технологичния процес



⁵ Web-scraping: Applications and Tools. (2015). European public sector information platform, Topics report № 10. https://www.europeandataportal.eu/sites/default/files/2015_web_scraping_applications_and_tools.pdf.

В етапа **Интернет достъп** се изпълняват две логически функционалности - „URL Searcher“ и „Scraper“: функционалността на блока **URL Searcher** е да намери и състави списък от уебсайтове, свързани с дадено предприятие. Обикновено този списък се получава чрез заявка в интернет търсачка, използвайки името на предприятието като „дума за търсене“. Основното допускане е, че ако предприятието има официален уебсайт, то той трябва да се намери в рамките на резултатите, генерирани от уебтърсачката. Блокът **Scraper** е отговорен за придобиване на наличното съдържание за всеки URL адрес в списъка с URL адреси, предоставени като вход. Той може да има допълнителни функции като филтриране на URL адреси (ако е предоставен списък с такива) и обикновено е конфигурируем чрез задаване на различни параметри като дефиниране ниво на извличане на данните (само началната страница или началната страница плюс първо ниво и т.н.).

Вторият етап **Съхранение** съдържа три функционалности „Retrieved URLs“, „Scraped content“ и „Index configuration“: блокът **Retrieved URLs** е основен контейнер с URL адреси (намерени като резултат от предишната стъпка), ранжирани от обикновен файл до система за управление на бази данни. Блокът **Scraped content** е контейнер със съдържание, получено от работата на блока „Scraper“. Обикновено е необходимо този блок да се реализира, прилагайки нетривиални решения, поради факта, че извлеченото количество информация е огромно и е съставено от неструктурирани данни. Блокът **Index configuration** представлява стратегия за индексирание на извлечените данни, съхранявани в блока „Scraped content“. В контекста на Big Data съдържанието, съхраненото огромно количество данни е параметър, който трябва да бъде взет под особено внимание, тъй като индексирането на данните дава възможност по-лесно да се открие информация в следващите фази.

Третият етап **Подготовка на данни** обхваща седем функционалности „Feature extraction“, „URL scorer“, „Tokenization“, „Data parsing“, „Word filters“, „Language specific lemmatization“ и „Term document matrix generation“. Блокът **Feature extraction** е отговорен за локализиране и намиране от извлечените от интернет данни на набор от предварително дефинирани характеристики на предприятията (например адреси, телефонни номера, имена, ДДС кодове и други). Обикновено функциите на този блок се изпълняват в специфична софтуерна програма. Блокът **URL scorer** се използва за оценяване на даден URL на базата на някои предварително дефинирани параметри, като наличието на някои характеристики, присъстващи в съдържанието на URL адреса. Като се има предвид, че в предишните стъпки е намерен определен списък с URL адреси, свързани с едно предприятие, този блок може да се използва самостоятелно или съвместно с друг блок, за да се идентифицира най-вероятният официален корпоративен URL адрес за конкретното предприятие. Блокът **Tokenization** обработва текстовото съдържание на

извлечените ресурси, като го трансформира в текст, който става вход за следващия етап, например като синтактичен анализ (parsing) и „text mining“ или за етапа „Анализ“. Блокът **Data parsing** се фокусира върху анализа на „символи“, произведени от блок „Tokenization“ чрез търсене на специфични регулярни изрази, съвпадащи изречения и т.н. Блокът **Word filters** се използва за филтриране на някои думи/символи (ако е предоставен списък от думи, които трябва да бъдат филтрирани) от „извлеченото“ текстово съдържание или за обогатяването на това съдържание със списък от водещи думи. Блокът **Language specific lemmatization** извършва т.нар. „лематизация“ на символите. В конкретния случай (компютърна лингвистика) лематизацията е алгоритмичен процес на определяне на лема на дума, основана на нейното първоначално значение. Когато не е възможно да се изведе първоначалното значение, обикновено се използва основната форма на символа, получен чрез прилагане на стриминг (автоматизиран процес, който произвежда основен стринг в опит да представя семантично свързани думи), който изчислява базовата форма на символа чрез работа върху една дума, без да е познато съдържанието. Блокът **Term document matrix generation** е отговорен за създаването на терминологично-документална матрица, която да бъде използвана от блоковете в следващия етап „Анализ“. Обикновено всяка клетка на матрицата съдържа броя на случаите на даден символ в уебсайта на предприятието.

В етапа **Анализ** функционалностите са три „Machine learning“, „Information extraction: NLP“ и „Information extraction: Deep learning“: Блокът **Machine learning** (и неговите подблокове) генерират статистически изходи чрез използване на класификатор или „обучител“ в контекста на компютърната лингвистика. Блокът **Information extraction: NLP** генерира статистически изходи чрез използване на подходи за компютърна обработка на естествен език (NLP). Блокът **Information extraction: Deep learning** генерира статистически изходи чрез използване на техники за дълбоко изучаване в контекста на подходите на изкуствен интелект.

Ключов въпрос при приложението на техниките на „web-scraping“ е дали структурата на съдържанието на уебсайта е известна предварително и могат ли да се правят различни допускания за структурата на данните преди да се извърши същинското „извличане“ на информация, или това не е възможно. Първият случай предполага използването на *специфичен „web-scraping“*, а вторият - *генеричен „web-scraping“*.⁶

Специфичен „web-scraping“ се прилага, когато са добре познати както структурата, така и (типът) съдържанието на уебсайтовете, от които ще бъдат извлечени данни. В този случай

⁶ Boeing, G., Waddell, P. (2016). New Insights into Rental Housing Markets across the United States: Web Scraping and Analyzing Craigslist Rental Listings. Journal of Planning Education and Research, 23 august. <http://journals.sagepub.com/doi/pdf/10.1177/0739456X16664789>; Vargiu, E., Urru, M. (2013). Exploiting web scraping in a collaborative filtering- based approach to web advertising. Artificial Intelligence Research, vol. 2, № 1. <http://www.sciedu.ca/journal/index.php/air/article/view/1390>.



компютърните работи възпроизвеждат човешкото поведение, като посещават автоматично уебсайта на дадена фирма с цел събиране на информация, която представлява изследователски интерес. Типични примери за такъв вид „web-scraping“ могат да се открият в областта на статистиката на цените, където повечето от позициите в онлайн магазина имат един и същ продуктово списък или продуктова страница. По този начин софтуерът за „извличане“ може да открива специфични елементи като: първоначална цена, продажна цена, етикет на стоката, описание, количество, цвят, размер и т.н. от много на брой уебстраници за много продукти.

Генеричен „web-scraping“ се прилага, когато липсват априори познания за структурата и съдържанието на фирмения уебсайт и е необходимо цялото му съдържание да бъде извлечено и обработено, за да се събере някаква информация, която представлява интерес за изследователя. Типичен пример за такъв вид „web-scraping“ е намирането на уебстраници на предприятията за извличане на някои общи характеристики на съвкупността. В този случай е необходимо да се разработи по-обща методология за „web-scraping“, както и за прилагане на специализиран софтуер за извличане и обработка на данните.

Разграничението и съпоставянето на двата вида „web-scraping“ е необходимо, тъй като при тяхното практическо приложение има технически и методологически разлики. Колкото по-малко се познава структурата на обекта, за който ще се прилага „web-scraping“, толкова по-обща са технологичните и методологичните подходи, които ще се използват на практика.

От технологична гледна точка със *специфичния „web-scraping“* е възможно предварително да се проучи обектът (уебсайтове или страници) и да се проектира специфичен „софтуер-скрапер“, който използва това знание, за да се навигира в сайта чрез идентификатори на структурата на страниците като например „html id“, „xpath“ и „css“ селектори. С *генеричния „web-scraping“* обикновено се отстраняват всички html маркировки и се прилага софтуер за извличане на съдържание от текст и документи върху останалото съдържание.

От методологична гледна точка със *специфичния „web-scraping“* се постигат добре дефинирани променливи, които да се използват при обработката на данните. Трябва да се има предвид, че при извличането винаги има известна степен на несигурност по отношение на емпиричните данни. При *генеричния „web-scraping“* обикновено се прилагат машинно самообучителни методи върху резултатите от стъпките за обработката на данни.

Техниката „web-scraping“ се използва от началото на съществуването на интернет, но търпи развитие през последните години. Прокси сървърите извличат съдържанието на уебсайтове и ги съхраняват за обслужване на локални компютри преди повече от 20 години. Тази дълга история на „web-scraping“ е довела до наличието на различни инструменти и методи, които могат да бъдат използвани за събиране на информация от уебсайтове, както и техники за анализ на

извлечената информация. Независимо от тяхното разнообразие те могат да се обединят в две групи: **подходи за машинно самообучение (machine learning) и детерминистични подходи.**⁷ За да се обоснове изборът на съответните методи за приложението им в пилотните проекти за дефинираните „use-cases“, е необходимо да се изясни накратко тяхната същност.

В контекста на официалната статистика извличането на данни *чрез „machine learning“ подходи* се състои в прилагане на алгоритми или модели, производни от набор от данни за самообучение, които се предполага, че са представителни за даден изучаван проблем. Параметрите на модела обикновено се настройват с набор за валидиране, преди да се измери неговата ефективност на т.нар. тестови набор от данни с известни характеристики. Накрая моделът се прилага към други набори от данни, за които не е известно нищо, но за които се предполага, че моделът работи добре, за да се произвеждат статистически данни. От друга страна, извличането на данни за официалната статистика *чрез детерминистични подходи* се състои в прилагане на алгоритми, проектирани от набор от правила с известни предварително характеристики на уебсайтовете и структурата на данните. Казано по друг начин - знанието на експерта се използва за проектиране на алгоритъм за обработка и интерпретиране на входни данни от уебпространството и други източници в статистически цели променливи. Този метод се нарича детерминистичен, защото алгоритъмът, приложен към едни и същи данни, винаги ще има един и същ (детерминистичен) резултат в сравнение с machine learning подхода, който зависи в голяма степен от обучителния набор от данни, който е използван.

Като цяло превръщането на данните от интернет в статистически данни обикновено изисква много стъпки, като във всяка стъпка може да се приложи различен подход. Един от факторите, които влияят върху избора, е сложността на взаимовръзката между входните променливи или производните характеристики, получени от входните данни и статистическите цели променливи. Ако тази връзка е сравнително ясна, детерминистичният подход е подходящото решение и обратно, ако тази връзка е комплексна, неизвестна или трудна за моделиране в алгоритъм, което обичайно се случва при работа с уебданни, подходът за машинно самообучение е правилният избор.

Важно е да се отбележи, че при подходите за машинно самообучение наличието на набор от данни за обучение със задоволително качество е от съществено значение. В много случаи това е предизвикателство. В някои от пилотните проекти тези данни за обучение са налични или могат да бъдат получени от предишни статистически изследвания. Това би могло да е валидно в

⁷ Alpaydm, E. (2004). Introduction to Machine Learning (Adaptive Computation and Machine Learning). MIT Press, ISBN 0-262-01211-1; Witten, I., Frank, E., Hall, M. (2011). Data Mining: Practical machine learning tools and techniques. 3rd ed. Morgan Kaufmann. ISBN-13: 978-0123748560.

краткосрочен план, когато в официалната статистика е въведен machine learning подходът, с цел да замени (частично) традиционния статистически процес, но в дългосрочен план това почти не е възможно. Очевидно е, че детерминистичните подходи нямат това предизвикателство, но имат други проблемни въпроси, които ги правят не толкова универсално приложими. При подхода на машинното самообучение от решаващо значение е превръщането на текст, събран чрез генеричен „web-scraping“ в модел. HTML и друга информация, съдържаща се в таговете и изображенията, е неструктурирана и съдържа много шум, който, ако не е филтриран, би направил сигналите неразбираеми. Поради тази причина трябва да се прилагат едновременно техники за извличане на съдържание от текст и данни, за да се структурират и стандартизират данните и да се открива съответната информация.

Какви методи и техники са използвани при реализацията на „use-cases“, дефинирани в настоящото емпирично изследване? В по-голямата си част събраните текстове са обработени чрез последователно изпълнение на следните стъпки: а) нормализиране: стеминг и лематизация; б) подбор на характеристики. За всяка зависима променлива (електронна търговия, присъствие в социалните медии) всички нормализирани термини са обработени, за да се открие най-подходящият термин за прогнозиране. Използвани са редица техники, като кореспондиращ анализ, регуляризационни техники (LASSO, Ridge, Elastic Net) - част от алгоритмите за машинно обучение. Целта е да се намали броят на термините до управляем брой значими независими променливи, които са използвани като вход към модели, пасващи в дадени обучителни набори от данни (получени по различни начини: чрез ad-hoc анализ на редица случаи или чрез използване на данни от регулярното статистическо изследване „Използване на ИКТ от предприятията“). По принцип обучителният набор от данни е разделен на отделни водещи елементи, като тестови набори от данни, които се използват за оценка на модела чрез сравняване на наблюдаваните и прогнозираните стойности.

Първият „use-case“ „Генериране на списък с URL адреси на предприятията“ се изпълнява на три стъпки: получаване на изходни точки от интернет търсачка; извършване на web-scraping на намерените URL адреси или друга извлечена от уебмрежата информация за тях; определяне кой от получените URL адреси като резултати от предходните две стъпки е истинският уебсайт на дадено предприятие. Първата стъпка се извършва или чрез структуриран приложно-програмен интерфейс API, или чрез извличане на уебстраница като резултат от търсещата машина. И двата случая са характерни примери за *специфичен* „web-scraping“ с *прилагане на техники за детерминистичен анализ*, т.е. резултатите от търсенето могат да се разглеждат като структурирано съдържание. Втората стъпка е пример за *генеричен* „web-scraping“, тъй като предварително не е известно нищо за обекта, който ще бъде подложен на

изследване. При третата стъпка - фазата на анализа, се използват подходи за машинно самообучение, за да се определи коректният URL адрес от всички намерени URL адреси, както и ръчно валидиране на резултатите.

Вторият „use-case“ „Електронна търговия в предприятията“ е по същество *генеричен „web-scraping“*, при който изходната точка е списък с фирмени URL адреси, за които предварително не е известно нищо за структурата на уебсайта. Използват се подходите на *машинно самообучение* за идентифициране на електронната търговия, като едновременно с това изследването „Използване на ИКТ от предприятията“ се използва като основно за набор от данни за обучение.

Третият „use-case“ „Присъствие на предприятията в социални медии“ се реализира чрез *генеричен „web-scraping“*, последван от *машинно самообучителен* подход. Резултатът е списък с характеристики на социалните медии, като за сравнителен анализ се използва статистическото изследване „Използване на ИКТ от предприятията“.

Четвъртият „use-case“ „Генериране на списък с URL адреси на предприятията“ чрез прилагане на италиански софтуер (разработен от ISTAT) е аналогичен на първия чрез използване на същите методи.

Не по-малко важен е въпросът за **програмните езици, ИТ инструментите и библиотеките**, използвани в пилотните „use-cases“. ИТ средствата за извършване на „web-scraping“, както вече беше отбелязано, биват генерични и специфични. Типичните генерични инструменти са следните: „import.io“, „Scrapy“, „imacros“, „Apache Nutch“ и други подобни. Втората група инструменти включва библиотеки за конкретни цели, като програмите „Tweepry за Python“, използвани за извличане на данните от Twitter. Подробен списък с web-scraping инструменти и библиотеки е достъпен в различни хранилища, включително в Github. Благодарение на разнообразието от инструменти и методи за „web-scraping“ всеки пилотен проект можеше да бъде реализиран по различен начин. Една от целите на текущото емпирично изследване от технологична гледна точка беше да се избере популярен, с отворен код или безплатен софтуер, който служителите на НСИ познават. Това е основната причина да се използва традиционен, както и специфичен софтуер за Big Data, който може да бъде изтеглен и приложен в НСИ без допълнителни разходи.

Целта на избора на инструменти за средства за съхранение е да се създаде среда, която да се поддържа лесно и CSV файловете да са с възможност за общ избор. Най-разпространеният начин за съхраняване на данни е файлова система от типа CSV файл. Дефинираните пилоти използват MySQL (релационни бази данни) и NoSQL Apache Solr (нерелационни бази данни). Решението за използване на файловата система като първично съхранение на данни е резултат от

факта, че повечето от инструментите, използвани в пилотите, имат вградени библиотеки за достъп до CSV файлове. От друга страна, CSV файловете се използват за съхраняване на резултатите от анализа. Използването на този тип файл позволява зареждането на данните в популярни приложения като R или MS Excel. За да се увеличи производителността на достъп до такива файлове, има възможност те да се съхраняват в HDFS файлова система (Hadoop Distributed File System), за да се извърши автоматичен и много ефективен паралелен достъп до данните. Целта на използването на Apache Solr в „use-case“ № 4 е да се осигури мащабируема среда, която да може да съхранява различни видове данни. Въпреки това основната цел на използването на Apache Solr е да съхранява уебсайтовете на предприятията в NoSQL база данни. Този тип база данни позволява динамично търсене чрез нейното съхранение, включително пълно текстово търсене, подчертаване на удари, фасетирано търсене, динамично клъстеризиране, интегриране на базата данни, обработка на богати документи, разпределено търсене, репликация на индекси и висока скалируемост.

За целите на настоящото емпирично изследване бяха разработени няколко различни софтуерни инструмента. Например ISTAT е разработил софтуер за URLSearcher в Java, който позволява да се извличат URL адреси за различни предприятия въз основа на техните атрибути, като име, град, адрес за връзка. Този софтуер е приложим за изпълнението на „use-case“ № 4 върху набор от данни на НСИ, тъй като може да се използва в Java среда във всяка операционна система. Резултатът е CSV файл или набор от файлове, включително идентифицирани връзки. По-нататъшен анализ на URL адресите се извършва с други потребителски Java приложения за търсене на URL адреси: URLSearcher, RootJuice и UrlScorer. Софтуерът е разработен като кратки скриптове, използвани за конкретни действия относно събирането и обработката на данни. Най-популярната среда за тестване на пилотите е използването на виртуална машина (например Linux Mint) или специална машина (например Linux Ubuntu Server, MS Windows).

Придобитият опит по време на изпълнението на четирите пилотни проекта показва, че софтуерните скриптове и приложения са независими от ИТ платформата. Например възможно е да се стартират Python скриптове или Java приложения в Linux, както и в Windows среда без никакви промени в изходния код. Ефективността на изпълнението е ключов проблем, особено когато се изтеглят и обработват целите уебсайтове. Обработката на неструктурирана информация изисква голям обем на централния процесор и памет, особено при прилагане на алгоритми за машинно обучение и като резултат обработката на данните не е много ефективна. Поради факта, че по-голямата част от средата, използвана за „сценариите“, има ограничени ресурси на

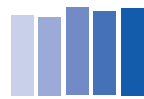
процесора и паметта, е трудно да се прецени колко ефикасен ще бъде алгоритъмът в реална производствена среда.

Въз основа на обработката на данните на НСИ със специално разработения за целта софтуер може да се направи заключение, че конвенционалните ИТ инструменти са достатъчни за създаването на списък с URL адреси за уебсайтовете на няколко десетки хиляди предприятия. Софтуерът, разработен и използван за внедряването на пилотите, е безплатен и с отворен код макар някои негови компоненти да са разработени „по поръчка“ за специфични цели. Това означава, че всеки може лесно да тества и подобри всеки от инструментите. От друга страна, при използването на такива софтуерни средства невинаги можем да разчитаме на добра документация или на подробно ръководство, за да се направи всичко работещо.

Въз основа на прилагането на италианския софтуер за масива от данни на НСИ може да се твърди, че Apache Solr има технически проблеми по отношение на Solr Connection pool в етапа на зареждане, веднага след фазата на „извличане“ на данните. Обстоятелството, че ИТ средствата за обработка на Big Data се променят много често, а паралелно с това - и технологията на уебсайтовете, е необходимо да се осигури възможно най-гъвкаво разработване на този тип ИТ инструменти. Това поставя въпроса за устойчивостта на използваните технологии. ИТ инструментите за пилотите вероятно ще се променят през следващите няколко години и затова не се препоръчва използване на конкретен компютърен език, за да е възможно лесно да се премине към други платформи.

Въз основа на българския опит изискванията за съхранението на база данни за около 27 000 предприятия отнемат памет около 1 GB твърд диск, включително данни от бизнес регистъра, извлечени данни чрез търсене на API, уебсайтове на предприятията и електронни магазини от заглавия на първа страница, ключови думи, описания и данни от URL адреси. Това позволява да се направи заключение, че за тези конкретни „сценарии“ е възможно да се използват традиционни технологии. Както вече беше споменато, възможен е избор между файлова система (CSV, JSON и т.н.), NoSQL база данни (Solr, Cassandra, Hbase и т.н.) или релационна база данни (MySQL, PostgreSQL, SQL Server и други). Решението за използването на конкретно средство за съхранение на информацията трябва да се вземе в зависимост от обема и вида на данните, които ще бъдат съхранявани. В този контекст е и въпросът с дублирането на данните. Необходима е рамка за премахване на дублирането, която автоматично ще изключи всички дублирания на уебсайтове и конкретна информация, взета от тях.

Като възможен резултат от разработените пилотни проекти са изчислени някои допълнителни индикатори като: относителен дял на извлечените URL адреси от списъка на



предприятията, относителен дял на предприятията, ангажирани с електронната търговия чрез уебсайтовете на предприятията, относителен дял на предприятията, които присъстват в социалните медии на уебсайтовете на предприятията, брой заети и други показатели на ниво област. За съжаление, все още тези резултати са достъпни само чрез Интранет мрежата на НСИ и не са публично достъпни. Прогнозните стойности, получени от източниците на Big Data, могат да се използват за постигане на две цели: на ниво единица - за обогатяване на информацията, съдържаща се в регистъра на изучаваната съвкупност, и за получаване на оценки на ниво съвкупност. Качеството на данните на ниво единица може да бъде измерено, като се вземат предвид същите показатели, произведени за оценка на модела, пасващ на обучителния набор от данни. При определени условия (ако обучителния набор от данни е представителен за цялата съвкупност) измерването на точността, както и на чувствителността и специфичността, изчислени в тази подгрупа, може да се счита за добра оценка на качеството на получените резултати от събраните Big Data.

Обратно, измерването на качеството на оценките на съвкупността, за които се използват прогнозни стойности, е много по-сложно. Пълното измерване трябва да се основава на оценка на средната квадратична грешка, т.е. съвместно отчитане на вариацията и изместването, влияещи на оценките на параметрите на съвкупността. Съществуват случаи, при които е възможно изчисляването на вариацията на оценителя, като се прилагат методи за повторно проектиране на извадката (по-специално bootstrap), вместо да се оценява изместването, изискващо познаване на истинската стойност на параметъра в съвкупността, което е рядко срещано условие на практика. Симулационните проучвания, при които се генерира изкуствена съвкупност, споделяща разпределителни характеристики с истинската съвкупност, могат да подпомагат оценяването на всички компоненти на средната грешка. Например в „сценариите“ за електронната търговия и представянето в социалните медии относителният дял на предприятията, чиито уебсайтове се характеризират с положителен отговор на тези променливи, може да бъде оценен в различни области чрез данни от традиционни статистически изследвания и чрез използване на интернет данни (с подход, базиран на модели).

Систематичната грешка при подбора на единиците, произтичаща от идентификацията на уебсайта, може да бъде особено важна - методите за идентифициране на URL адреси са по-склонни да работят добре за уебсайтове, които извършват електронна търговия. Двата набора от оценки могат да бъдат сравнени. За да се реши дали тяхната разлика по отношение на качеството е релевантна, сравнението може да използва техники за повторно проектиране на извадката и симулационни проучвания.

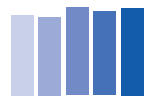
II. Практическа реализация на пилотните „сценарии“

1. Use-case 1: Генериране на списък с URL адреси на предприятията (URLs retrieval)

В пилотния проект е използвана информация от статистическия Бизнес регистър на НСИ, предприятията от който формират съвкупността на изследването „Използване на ИКТ от предприятията“. Основната цел се състои в генериране на списък от валидирани URL адреси на предприятията, който да бъде използван впоследствие за извличане на информация от корпоративните сайтове за наличие на електронна търговия и присъствие в социалните медии („use-case 2“ и „use-case 3“) чрез прилагане на различни ИТ техники. Екипът за реализацията на този „сценарий“ включва: ИТ специалист/програмист - да извършва дейностите по извличане на URL адреси от наличните набори с данни и да разработи софтуерни програми за прилагане на техниката „web-scraping; експерт, отговорен за поддържане на актуален списък с URL с адреси на предприятията; статистик, за да анализира концептуалното свързване между уебинформацията и бизнес регистъра, като решава възникнали методологични проблеми - например липсващи стойности на уебсайта на предприятието и др., и технически персонал. Работният процес за настоящия „use-case“ се състои от следните последователни действия: създаване на тестови набор от предприятия със свързани URL адреси от посочената целева съвкупност; използване на приложен програмен интерфейс API в мрежата за търсене името на предприятието или името на предприятието, последвано от „контакт“, както и съхраняване на първите 10 резултата като „кандидати“-уебсайтове.

За целите на настоящия „сценарий“ е разработен специален софтуер, като се използва определен брой API търсения от интернет търсачките - *Jabse* и *Google Search*. За всеки „кандидат“- уебсайт се оползотворяват събраните уебданни, за да се получи подробна информация. Това могат да бъдат извлечени данни от уебсайта на компанията или фрагмент от резултата от API за търсене. Използването на събраните данни за идентифициране на уебсайтове се извършва ръчно или чрез прилагане на автоматичен алгоритъм. В процеса на работа по този пилотен проект се достигна до следните няколко извода:

- Чрез прилагането на основната методология за изпълняване на заявки за търсене (API Search) на уебсайта на предприятието е възможно да се идентифицират голям брой уебсайтове на същото предприятие и след това да се осъществи свързване на уебданните с неговите характеристики.
- Възможно е да се идентифицират както фалшиви положителни данни (неправилно идентифицирани уебсайтове), така и фалшиви негативни данни (уебсайтове, които не са



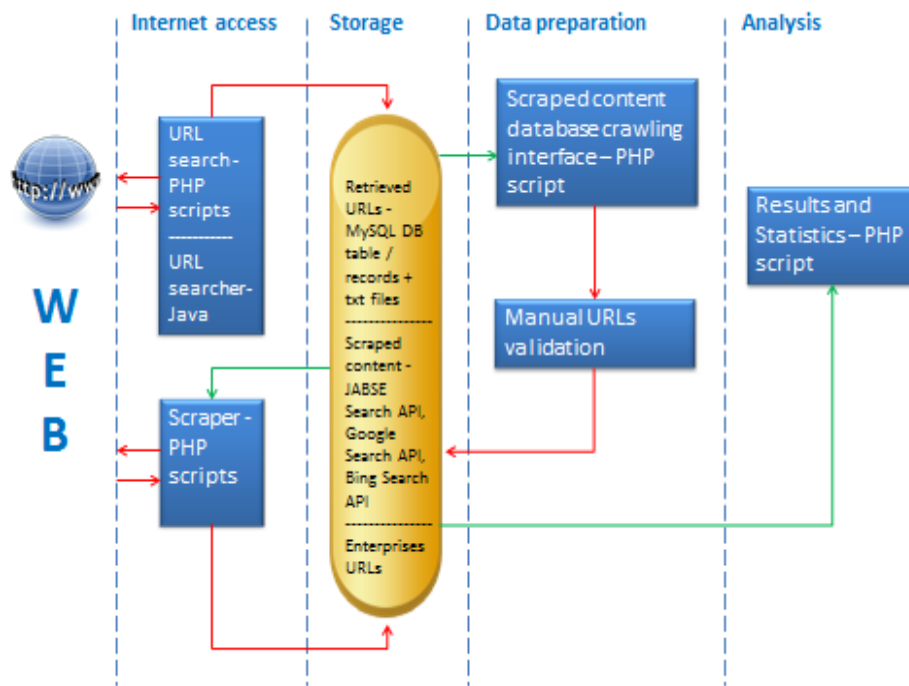
идентифицирани). За проверка на „гранични“ случаи се изисква ръчна проверка, извършена от статистически експерти - особено при значително малък набор от данни. За да се създаде списък на URL адреси на предприятията с добра точност, винаги ще е необходима известна „експертна“ помощ.

- Разнообразието от уебданни може да е полезно при избора между „кандидат“ уебсайтове API Search - включително данни, извлечени директно от уебсайтовете, „фрагменти“ от резултатите от търсенето, класиране на резултатите от търсенето и информация за регистрация на уебсайтове. Обаче трябва да се има предвид, че не е достатъчно да се използва само регистрационна информация на предприятието, което предполага, че някои данни трябва да бъдат събрани от уебсайтове. Това означава, че всяка официална статистическа институция, която се интересува от идентифициране на уебсайтове на предприятия, трябва предварително да е разрешила всички правни и етични въпроси, произтичащи от националното законодателство по отношение на прилагането на техниката „web-scraping“.

- Когато методът за идентифициране на уебсайтове не може да идентифицира всички уебсайтове на предприятието, е важно да се отчете стеминг отклонението от вероятността да бъде идентифициран всеки от дадените уебсайтове.

Нека си припомним, че целевата съвкупност, обект на „web-scraping“, се състои от 26 836 предприятия, 20 649 имейла и 2 006 URL адреса. Използваните подходи са основно два: приложение на Jabse Search API и Google Custom Search API на базата на името на предприятието с филтриране на резултатите от Search API и извличане на URL адреси от набори от данни, съдържащи Булстат номер, URL адреси, данни за контакт и други характеристики на предприятието от статистическия Бизнес регистър. Обща представа за производствения поток може да се получи от фиг. 2, на която е представена логическата архитектура на Use-case 1.

Фиг. 2. Логическа архитектура на Use-case 1: URLs retrieval



Компонентът „URL Searcher“ използва информация за URL адреси и имейли от бизнес регистъра, за да проверява, верифицира и генерира имена на домейни и да обработи списъка с URL адреси. Средството Scraper използва фирмените имена и чрез ефективната работа на интернет търсачките Jabse Search API и Google Custom Search API се събира набор от 10 предложени/вероятни URL адреси на дадено предприятие от съвкупността. Събраната информация от URL Searcher и Scraper се съхранява в специално създадена за целта база данни. Истинските URL адреси на предприятията се идентифицират чрез софтуерен интерфейс (web-crawling), специално разработен за „обхождане“ на базата данни, които статистическите експерти на последната фаза от процеса валидират ръчно. Във фазата на анализ статистическите резултати са изчислени със специфичен софтуерен скрипт. След като компонентът URL Searcher получи списък с предприятия с налични URL адреси и имейли, същият проверява дали първоначалните URL адреси са реални уебсайтове и съхранява резултатите в базата данни. Ако URL адресите не са потвърдени или липсват, то тогава URL адресите се генерират от имената на домейните на електронната поща, като се изключват популярни домейн наименования на имейлите (като gmail, yahoo и др.), в случай че са налични. Генерираните URL адреси се верифицират за действителното им съществуване от Searcher. Всички проверени и установени като действителни URL адреси са съхранени в базата данни (**7 038 URL адреса**).

Компонентът Scraper използва автоматизирания интерфейс за търсене на *Jabse*, като получава до 10 резултата от търсенето за наименованието на дадено предприятие на български език и до 10 резултата от търсенето за наименованието на предприятието, транслирано на латиница. Тогава Scraper изключва от резултатите от търсенето сложните URL адреси и записва само тези, които са близки до наименованието на домейна, като предполага, че това са най-вероятните адреси. Резултатите се **съхраняват в базата данни в текстов и html формат (15 638 серии** с до 10 най-вероятни резултата от търсенето на български език, **16 201 серии** с до 10 най-вероятни резултата от търсенето на латиница). След това Scraper използва интерфейса за търсене на Google, като получава до 10 резултата от търсенето за наименованието на фирмата на български език и записва резултатите в базата данни в **json формат (26 829 серии** с до 10 резултата от търсенето). Интерфейсът за „обхождане“ на базата данни е използван като помощно средство от статистическите експерти, за да избират реалните URL адреси на фирмите от всички предложени/намерени URL адреси от URL Searcher и Scraper. Резултатите от тази фаза са **9 809 реални и действителни URL адреса**. Скриптът за резултати и статистики дава информация за корпоративните URL адреси в реално време.

Що се отнася до технологичния избор към момента на изпълнение на този пилотен проект („use-case 1“) НСИ нямаше никакъв опит в областта Big Data и по отношение на прилагането на техниката „web-scraping“. Поради тази причина първият избор на технологични инструменти са безплатни софтуерни уебприложения, с които само един експерт в НСИ имаше известен опит: Apache уебсървър, MySQL база данни и PHP програмен език. В процеса на работа са използвани PHP за програмния софтуер, MySQL за платформата за съхранение и Apache за изпълнение на PHP скриптове през уеббраузърите. В специално разработения по поръчка софтуер са интегрирани и използвани предложените/намерените резултати от уебприложенията *Jabse Search API* и *Google Custom Search API*. PHP скриптовете бяха изпълнени в браузър с използване на метамаркер за опресняване на HTML съдържание (например скриптът задава заявки за приложния програмен интерфейс (Search API) на всеки три секунди с фирмени данни и съхранява информацията, събирайки я в базата данни.

След успешното завършване на практическата реализация, могат да се направят няколко **заклучения:**

- По отношение на методологията може да се твърди, че *Google Search API* дава най-добри резултати, като осигурява 200 търсения на ден безплатно или 1 000 търсения за сумата от 5 EUR до максимум 10 000 търсения на ден. От друга страна, базата данни на *Jabse* не обхваща всички предприятия. *Jabse Search* работи по-добре с английската си версия отколкото с българската, но за съжаление Search API покрива само българската версия. Като цяло **26 836**

записа бяха проверени ръчно от експерти (част от екипа на проекта) в рамките на 45 работни дни или средно проверени по 600 записа за всеки работен ден.

- Конвенционалните ИТ инструменти са достатъчни за създаването на списък с URL адреси на десетки хиляди предприятия. Размерът на базата данни, съдържаща около 27 000 предприятия, заема около 1 GB хардуерно пространство (данни от Бизнес регистъра, данни от търсачките Search API, уебсайтове на предприятия, електронни магазини от наименованията на първите страници на уебсайта, ключови думи, описания и URL адреси).

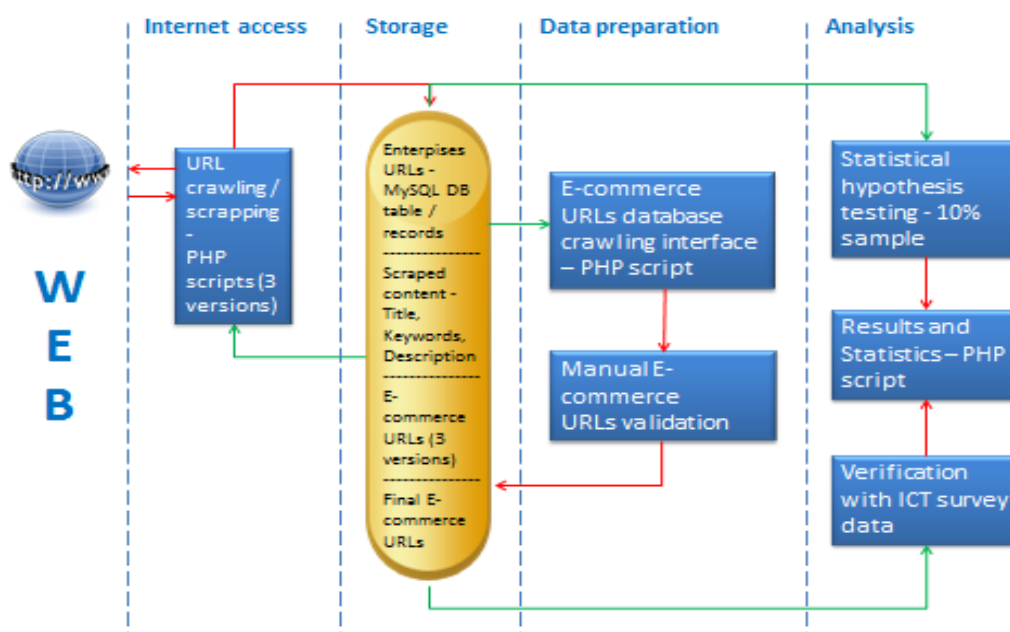
- Няма правни ограничения, защото са използвани публично достъпни ИТ инструменти и продукти (Jabse и Google Search APIs), за да се получат URL адресите на предприятията.

2. Use-case 2: Електронна търговия в предприятията (E-commerce)

За изпълнение на задачите по настоящия пилотен проект е използван списъкът с действителните URL адреси на предприятията (получени като резултат от работата по „use-case 1“), за да се установи дали дадено предприятие, което има фирмен уебсайт, се занимава с електронна търговия. Основната цел на този „use-case“ е прилагане на техниката „web-scraping“ върху заглавните уебстраници на официалните уебсайтове на фирмите и след това да може да се направи възможно най-точна прогноза за наличието на е-търговия на корпоративния сайт. Това става чрез използване на предварително дефинирана таксономия и набор от ключови думи, приложени към текстовото съдържание на уебсайта, получено чрез извличане на съдържание от текст. За да се постигнат по-качествени резултати, извличането на съдържание от текст е направено по три различни начина (три прогнозни версии). Получените крайни прогнозни резултати бяха валидирани ръчно, за да се оцени коя от трите прогнозни версии е най-точна. Очакваните резултати са дефинирани алтернативно: „успех“ - положителна или отрицателна оценка (дадено предприятие извършва или не извършва електронна търговия), и „неуспех“ - няма достъп до уебстраницата на предприятието. Работният процес за настоящия „use-case“ се състои от следните последователни действия: извличане на текстово съдържание от списъка с предварително дефинирани интернет страници на предприятията (от „use-case 1“) само на ниво заглавна страница на уебсайта на дадено предприятие. Обемът на извадката, от която се извлича текстово съдържание, е 9 809 уебсайта на предприятия; създаване на характеристики въз основа на наличието или отсъствието на думи в текстовото съдържание: създадени са списъци или „речници“ на ключови думи (няколко групи от подобни списъци), които индикират наличието или отсъствието на електронна търговия на сайта на дадено предприятие; използване на тези характеристики и някои алгоритми, за да се предскаже дали дадено предприятие е ангажирано в

електронната търговия, а именно прилагане на автоматичен филтър, базиран на създадените характеристики в предишния етап на работа. Като цяло опитът за идентифициране на сайтовете за електронна търговия се оказва успешен, но усилията бяха насочени към постигане на баланс между прецизност и повторемост на изпълнението на задачата. Поради тази причина възникна необходимостта от по-нататъшно разработване на специфични методи, преди да се стигне до получаване на надеждни оценки. Изпълнението на един по-опростен метод, основан на правила за идентифициране на създадените характеристики, може да се оцени като сравнително добро. Въпреки това ефективното прилагане на модел от типа „кошница с думи“, където всяка дума се третира независимо и използването на по-съвременни техники - като например от типа NLP техники, биха дали вероятно още по-добри резултати. Логическата архитектура на „use-case 2“ е представена на фиг. 3:

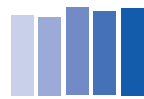
Фиг. 3. Логическа архитектура на Use-case 2: E-commerce



Средството *URL crawling-scrapping* използва информация от списъка с URL адресите, за да „посети“ уебстраниците на предприятията, да предскаже URL адресите за електронната търговия на предприятията и да извлече заглавие, ключови думи и описателно съдържание от първата страница на фирмения уебсайт или да предскаже първите страници на уебсайтове за електронна търговия. Събраната по този начин информация е съхранена в базата данни.

Интерфейсът за „обхождане“ на базата данни (*URLs database crawling interface*) за електронна търговия се прилага като помощно средство от статистически експерти-доброволци. Същите проверяват и валидират ръчно коректния URL адрес за електронна търговия на предприятието, използвайки информацията, събрана от средството *URL crawling-scraping*. Във фазата на анализа бяха приложени класическите методи за проверка на статистически хипотези, за да се провери колко точни и изчерпателни са използваните „web-scraping“ алгоритми за предсказване. В допълнение, окончателните резултати от този пилотен проект бяха сравнени и верифицирани с данните от традиционното статистическо изследване „Използване на ИКТ в предприятията“ за 2016 г., провеждано от НСИ. Статистическите резултати от сравнението са изчислени със специфичен софтуерен скрипт. Средството *URL crawling-scraping* взема уебадреса от списъка с URL адресите на предприятията и извлича съдържанието на първата страница на фирмения уебсайт. След това с помощта на PHP скрипт с три логически алгоритъма (използвайки 4 положителни и 1 негативен списък с ключови думи) прогнозира URL за електронната търговия на предприятието, ако тя съществува. Резултатите са, както следва: чрез алгоритъм № 1 са прогнозирани 1 139 URL адреса за е-търговия; чрез алгоритъм № 2 - 1 048 URL адреса, и чрез алгоритъм № 3 - 662 URL адреса за електронна търговия на предприятията. *URL crawling-scraping* компонента „извлича“ уебадреса на фирмата за електронна търговия, заглавието, ключовите думи и описанието на уебсайта и съхранява извлечената информация в SQL база данни заедно с прогнозираните вероятни URL адреси за електронна търговия. Интерфейсът за обхождане на базата данни за електронната търговия с URL адреси (*e-commerce URLs database crawling interface*) представя резултати от предвиждането за е-търговия, първите уебстраници на електронните магазини и заглавията на първите страници на предприятията, ключови думи и текстови описания. На следващ етап статистически експерти-доброволци валидираха получените резултати ръчно и само реалните URL адреси за е-търговия на предприятията бяха съхранени окончателно в базата данни. По този начин са открити общо **856 уебстраници за електронна търговия**.

Отрицателната оценка е тествана с прилагане на класическия метод за доказване на статистически хипотези (10% извадка) при допускане за 90% точност и 80% пълнота на прилагания алгоритъм. За целта беше излъчена 10% случайна извадка от двете подсъвкупности: е-търговци и останалите, които не са такива, като статистически се проверява дали точността е по-ниска от 90% и пълнотата е по-малка от 80%. Прилагайки нормалното разпределение за нулевата хипотеза, се стига до извода, че използваният филтър е точен и изчерпателен. В резултат на това беше установено, че едва 27 е-търговци не са обхванати от алгоритмите за прогнозиране. В допълнение, очакваният резултат от този пилотен проект беше да се вземе решение дали



произведената информация по този иновативен начин може да се използва за замяна на някои въпроси, съдържащи се във въпросника на традиционното изследване „Използване на ИКТ от предприятията“, с цел намаляване на тежестта на респондентите и получаване на реална, но по-качествена информация. Верификацията на резултатите е извършена с данните от изследването „Използване на ИКТ от предприятията“. След бенчмаркинг анализа между данните от традиционното изследване в областта на ИКТ и получената от настоящия пилотен проект информация, се получават следните резултати: от 26 836 предприятия (обхват на проекта) в извадката за 2016 г. на изследването попадат 4 332 предприятия от тях. Намерени са 89 нови предприятия, извършващи в действителност електронна търговия, които са включени в обхвата на традиционното изследване, но дават отрицателни отговори при попълване на въпросника от анкетното проучване. Като правило, ако е-магазините на предприятията имат основни уебсайтове или имат връзка към електронни магазини, то те в повечето случаи се откриват на първите страници от уебсайтовете на предприятията. В някои случаи специалистът, разработващ фирмен уебсайт, поставя връзка към собствения си електронен магазин, така че трябва да се използва негативен списък от ключови думи със създателите на уебсайтове. Трябва да се има предвид, че по-дългият списък с ключови думи невинаги дава по-добри резултати за разлика от варианта, когато по-стриктният алгоритъм дава по-точни резултати, но пък пропуска повече URL адреси за е-търговия.

След успешното завършване на практическата реализация, могат да се направят няколко **заклучения:**

- По-свободният алгоритъм намира повече уебсайтове, които в действителност не са електронни магазини. Ако от излъчената 10% извадка, са получени 856 уебстраници за електронна търговия и са открити 27 пропуснати е-търговци, то тогава се получават общо 1 126 вероятни URL адреса на фирми за е-търговия ($856 + 27 \times 10 = 1126$). Резултатът е близък до прогнозирания в началото брой от 1 139 URL адреса за е-търговия на предприятията, прилагайки по-свободния алгоритъм. Като заключение, общо 11.5% от предприятията, които имат уебсайтове, извършват е-търговия и 4.2% от общия брой предприятия в съвкупността извършват електронна търговия.
- Използваните ИТ средства и инструменти са достатъчно ефективни за изпълнението на настоящия пилотен проект.
- Не са възникнали правни ограничения по отношение на прилагането на техниката „web-scraping“.

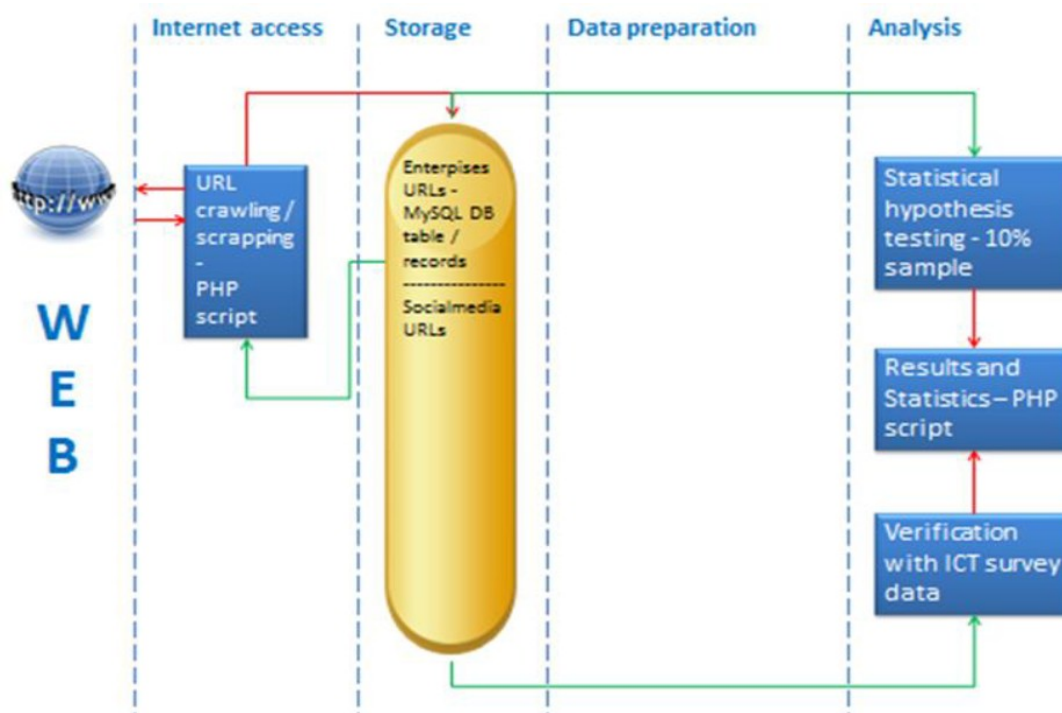
3. Use-case 3: Присъствие на предприятията в социални медии (Social media presence)

За изпълнение на дейностите се използва списъкът с URL адресите на предприятията (получен като резултат от изпълнението на „use-case 1“), за да се провери дали дадено предприятие присъства, или не в социалните медии (Facebook, Twitter, LinkedIn, Google, YouTube, Pinterest, Instagram). Основната цел на този пилотен проект е да предостави информация за активността на българските фирми в социалните медии, т.е. всички профили в социалните медии са взети под внимание. Общата концепция за този „use-case“ е да се извлече уебстраницата и после да се търсят връзки към профилите в социалните медии, включени в обхвата на този проект. В бъдеще е възможно да се добавят още атрибути за събиране на информация, например дали профилът е актуален и колко често се променя съдържанието му. Очакваният резултат от този пилот е вземане на решение дали данните за социалната медийна активност на предприятията могат да се използват за актуализиране на статистическия бизнес регистър и дали има засилено присъствие на българските компании в социалните медии. В допълнение, очакваните резултати могат да бъдат използвани за вземане на решение за замяна на някои въпроси, съдържащи се във въпросника на изследването „Използване на ИКТ от предприятията“. Работният процес за настоящия „use-case“ се състои от следните последователни действия: извличане на съдържанието на първите уебстраници с URL адреси от списъка с URL адреси на предприятията (получен като резултат от изпълнението на „use-case 1“); идентифициране на профилите на социални медии (предимно Facebook и Twitter) чрез проверка на връзките от извлечените страници и филтриране с наименованията на социалните медии; оценка на резултата чрез бенчмарк анализ с данните от традиционното статистическо изследване „Използване на ИКТ от предприятията“. Изводът, до който достигна екипът по време на работата по този „сценарий“, е: в повечето случаи профилите на предприятията в социалните медии са успешно идентифицирани. Въпреки това намерените профили чрез прилагане на „web-scraping“ са по-малко като обем от тези, получени като отговори на въпросника от изследването „Използване на ИКТ от предприятията“. Често срещан проблем е още, че идентифицираните профили в социалните медии не са непременно съответстващи на търсените URL адреси или предприятия. Не може да се твърди със сигурност дали фирмените профили в различните социални медии се актуализират, или не. Анализът на данните за социалните медии с Twitter API е ефективен, но възникват трудности с текущия Facebook API.

Екипът на проекта използва идентичен набор от ИТ средства за този пилотен проект, както за „use-case 1“ и „use-case 2“, а именно: *специално разработен софтуерен скрипт* за „извличане“ на уебсъдържанието от сайтовете на предприятията и подпомагане на експертите да

определят присъствието на дадено предприятие в социалните мрежи. Производственият поток на „use-case 3“ е представен на фиг. 4.

Фиг. 4. Логическа архитектура на Use-case 3: Social media presence



Средството „URL crawling-scrapping“ (същият инструмент е използван в „use-case 2“) използва информация от списъка с URL адресите на предприятията, за да „посети“ уебстраниците на предприятията и да прогнозира присъствието на предприятията в социалните медии. Информацията, събрана от този компонент, е съхранена в базата данни. Окончателните резултати са сравнени с данните от изследването „Използване на ИКТ от предприятията“ за 2016 година. Статистическите резултати са изчислени със специфичен софтуерен скрипт. Функционалното описание на компонента е следното: „URL crawling-scrapping“ „взема“ уебадреса от списъка с URL адреси на предприятията (9 809 URL адреса от „use-case 1“) и извлича съдържанието на първата страница на уебсайта на предприятието. След това, използвайки името на социалните медии, PHP скриптът търси уебленк към профила на социалните медии и съхранява намерената информация в SQL база данни. Получените резултати са, както следва: Facebook - 2 356 профила; Twitter - 922 профила; LinkedIn - 560 профила; Google - 871 профила; Youtube - 527 профила; Pinterest - 139 профила, и Instagram - 127 профила.

Установено е, че 24.9% от предприятията с уебсайтове имат поне един профил в социалните медии, а 9.1% от предприятията използват поне една от обхванатите в изследването социални медии. Получените резултати от „web-scraping“ са проверени за прецизност и изчерпателност чрез прилагане на методите за проверка на статистически хипотези. За целта беше излъчена 10% случайна извадка при гаранционна вероятност за 90% точност и 80% пълнота на прилагания алгоритъм. По-конкретно, беше тествано дали точността е по-ниска от 90% и пълнотата е по-малка от 80%. Прилагайки нормалното разпределение за нулевата хипотеза, се стига до извода, че използваният филтър е точен и изчерпателен. В резултат на това беше установено, **че само четири предприятия не са обхванати**. Верификацията на резултатите е извършена с данните от изследването „Използване на ИКТ от предприятията“. След бенчмаркинг анализа между данните от традиционното изследване в областта на ИКТ и получената от настоящия пилотен проект информация, се получават следните резултати: от 26 836 предприятия (обхват на проекта) в извадката за 2016 г. на изследването попадат 4 332 предприятия от тях. Намерени са 382 нови предприятия, които присъстват в социалните медии и попадат в обхвата на изследването.

След успешното завършване на практическата реализация, могат да се направят няколко **заключения**:

- По отношение на методологията съществува риск, че някои от връзките във Facebook или Twitter, представени на уебстраницата, може да са свързани с други предприятия. Поради тази причина е необходима оценка на Facebook и Twitter профилите, за да се предостави надеждна информация. В някои случаи предприятията могат да имат няколко Facebook профила. Поради тази причина е необходимо основният социален профил да се свърже с профила на компанията. Като правило връзките на профилите на социалните медии с профилите на предприятията се намират на първите страници на уебсайтовете на предприятията.

- По отношение на ИТ технологии е направен изводът, че използваните ИТ средства и инструменти са достатъчно ефективни за изпълнението на настоящия пилотен проект.

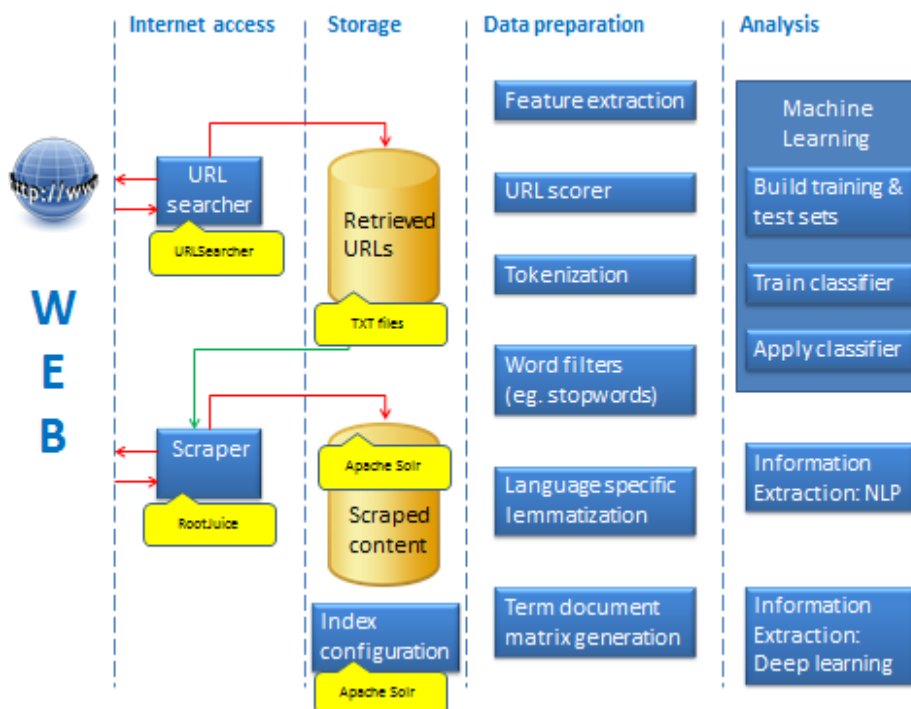
- И при този пилотен проект не са възникнали правни ограничения по отношение на прилагането на техниките „web-търсене“ и „web-scraping“.

4. Use-case 4: Генериране на списък с фирмени URL адреси на предприятията чрез прилагане на италианския софтуер (разработен от ISTAT)

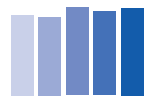
Целта на пилотния проект е тестване на софтуера на ISTAT за извличане на данни и инвентаризация на URL адреси с автоматизирана процедура за извличане на URL адреси в български условия. Базовата съвкупност е същата - предприятията в обхвата на изследването „Използване на

ИКТ от предприятията“. Задачите са аналогични с тези на „use-case 1“ - генериране на списък от URL адреси на предприятията, които да бъдат използвани впоследствие за „извличане“ на информация за е-търговия и присъствие в социалните мрежи. НСИ използва предложения софтуер с отворен код от ISTAT: Java Run time environment за програмите URLSearcher, RootJuice, URLScorer и URLMatchTableGenerator и Apache Solr платформа за съхранение. Решението за тестването на италианския софтуер върху българската съвкупност от предприятия беше взето с цел да се споделят добри практики и опит между европейските страни, участващи в проекта. Резултатите от този „use-case“ са използвани за бенчмаркинг с резултатите от „use-case 1“, като са направени оценки за постигнатите резултати в двата случая. Работният процес за настоящия „use-case“ се състои от следните последователни действия: използване на тестовия списък от предприятия със свързани URL адреси от „use-case 1“; използване на приложен програмен интерфейс (API) за търсене в мрежата на името на предприятието или името на предприятието, последвано от „контакт“ и съхраняване на първите 10 резултата като „кандидати“-уебсайтове. За целите на настоящия „сценарий“ е приложен италиански софтуер, като се използва определен брой API търсения от интернет търсачката Bing; за всеки „кандидат“-уебсайт се оползотворяват събраните уебданни, за да се получи подробна информация. Това могат да бъдат извлечени данни от уебсайта на компанията или фрагмент от резултата от API за търсене и използване на събраните данни за идентифициране на уебсайтове чрез прилагане на автоматичен алгоритъм или ръчно. Технологичната реализация на проекта е представена на фиг. 5.

Фиг. 5. Логическа архитектура на Use-case 4: Генериране на списък с фирмени URL адреси на предприятията чрез прилагане на италианския софтуер



Накратко, общото описание на производствения поток в логическата архитектура може да се представи по следния начин: средството URLSearcher използва фирмените имена с приложния програмен интерфейс API на Bing Search, за да може да се формира набор от 10 на брой предложени URL адреса и събраната информация се съхранява в txt файл. Компонентът RootJuice „взема“ txt файла, „извлича“ съдържанието на фирмените уебсайтове и записва информацията в csv файл. Информацията от csv файла се зарежда в Apache Solr (платформа за търсене на предприятия с отворен код), а UrlScorer използва съхранените вече данни в Apache Solr, за да генерира файл със зададени резултати за всеки от предложените URL адреси на предприятието. След това компонентът URLMatchTableGenerator получава резултатите от URLScorer и ги сравнява с предварително известен списък на URL адресите на предприятието. Получените резултати се анализират и се избира най-вероятният уебадрес на дадено предприятие. Във функционално отношение компонентът URLSearcher получава два файла: единият, съдържащ списъка с фирмените имена, и съответният списък с идентификационни номера на предприятията от бизнес регистъра (както отбелязахме 26 836 предприятия). За всяко предприятие Bing търсачката „извлича“ списък с първите 10 URL адреса, които се съхраняват във файл (всяка фирма има по един такъв файл). В края на тази дейност програмата чете всеки произведен файл и създава сийд файл във формат txt, съдържащ всички резултати. Програмата RootJuice приема като вход 3 файла: сийд файл от URLSearcher, списък с домейни на URL адреси, които трябва да се избягват (обикновено домейни на директории, жълти страници и т.н.) и □ конфигуриращ файл. Програмата RootJuice се опитва да достигне до HTML страница за всеки ред от сийд файла (ако URL адресът не е в списъка на домейните, които трябва да се избягват). От всяка достигната HTML страница програмата избира само текстовото съдържание на полетата, които представляват интерес, и ги записва като отделен ред CSV файл. След това CSV файлът от RootJuice се импортира в платформата за съхранение с отворен код Apache Solr. UrlScorer е програма, която чете един по един всички документи, съдържащи се в определена Solr колекция, и определя оценки на всеки от тези документи въз основа на стойностите на някои показатели. По-специално, изчислява се стойността на двоичните показатели, например: URL адресът съдържа деноминацията (Да/Не); „извлеченият“ уебсайт съдържа географска информация, съвпадаща с вече наличната в бизнес регистъра (Да/Не); „извлеченият“ уебсайт съдържа същия фискален код в бизнес регистъра (Да/Не); „извлеченият“ уебсайт съдържа същия телефонен номер в регистъра (Да/Не) и т.н. URL Match TableGenerator получава резултатите от URLScorer и ги сравнява с предварително известен списък от URL адреси на предприятието. Резултатът показва, че софтуерът прогнозира правилните URL адреси на 67% от общия брой предприятия, като има възможност за подобрене чрез адаптиране на по-добър списък с жълти страници и



интернет каталози, които трябва да отпаднат предварително. Също така трябва да се има предвид, че съществуват разлики между очакваните полета за данни от италианския софтуер и полетата на файла от бизнес регистъра на НСИ - например кодът на областта и телефонният номер на предприятието са свързани в българския регистър за разлика от тяхното отделно използване в софтуера на ISTAT. Тази ситуация също допринася за сравнително ниския процент на съвпадение на резултатите вследствие на прилагането на българския и италианския софтуер.

След успешното завършване на практическата реализация, могат да се направят няколко **заклучения:**

- По отношение на методологията се налага изводът, че добрият списък от бизнес каталози е основа за по-добри прогнози. Използването на точен брой полета с данни, необходими на софтуера, е основа за по-добро оценяване на предложените URL адреси.
- За ИТ технологиите може да се отчете, че използваната от ISTAT версия на Apache Solr е най-добрата, за да работи софтуерът подходящо, защото с последната актуална версия 6.5 на Apache Solr възникват технически проблеми.
- Констатирано е, че няма правни ограничения, защото са използвани публично достъпни ИТ инструменти и продукти (Bing Search APIs), за да се получат URL адресите на предприятията.

Заклучение

Проведеното емпирично изследване „Извличане на информация от интернет за характеристики на предприятията (web-scraping)“ е първо в Big Data практиката на Националния статистически институт. Приложените методи и техники за набиране, структуриране, обработка и анализ на тази информация предопределя неговата уникалност. Като основен резултат от работата по четирите „сценария“ в контекста на SGA-I се открояват няколко ключови бъдещи предизвикателства, свързани с прилагането на техниката „web-scraping“ за извличане на характеристики на предприятията от техните уебсайтове и текущите дейности на екипа относно работата по SGA-II.

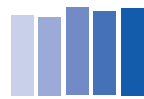
На **първо място**, отклоненията в крайните резултати са релевантни на стандартните отклонения в статистическата теория. Например при изпълнение на „use-case 1“ за автоматично намиране на URL адреса на предприятията от съвкупността стана ясно, че някои от фирмите са по-малко склонни да имат корпоративен уебсайт от други. Друг пример за отклонение е работата по „use-case 2“, където се оказва, че уебпаяка има способността да надценява идентифицирането на е-търговци и ако тези данни се използват директно за производство на официална статистика,

може да има значително отклонение от действителността, без да могат данните да бъдат ажустирани. Освен това се установи, че при някои уебсайтове, които основно използват JavaScript, възниква технически проблем, при който е трудно да се извличат данни чрез „web-scraping“ и това също води до непълнота на данните. Основно предизвикателство в тези случаи ще бъде да се анализират отклоненията и да се разработи метод, чрез който събраните данни от интернет да се коригират по начин, който позволява тяхното използване за оценки, класифицирани като официална статистика. Това вероятно ще изисква методологическа работа около комбинирането на извлечени данни, данни от статистически изследвания или административни данни.

На **второ място**, предизвикателства съществуват и по отношение на етичните въпроси. Все повече физически лица и предприятия са достъпни в онлайн форма чрез интернет източници. Това е значителна възможност, но и предизвикателство, което е вероятно да доведе до по-голяма загриженост сред обществеността относно начина, по който правителството и държавните институции използват онлайн данните. В конкретното изследване НСИ извлича публични текстови данни от фирмените уебсайтове, за да идентифицира е-търговците или присъствието на дадено предприятие в социалните мрежи. В отговор на това НСИ и Евростат ще имат нужда да разработят прозрачни политики за прилагане на „web-scraping“, за да смекчат или изобщо да отпаднат обществените опасения относно събирането и използването на тези Big Data.

На **трето място**, използваните методи за „web-scraping“ в настоящото емпирично изследване са изцяло базирани на методи за събиране на текстова информация от уебстраници в интернет пространството. Това може да се превърне в голямо предизвикателство, тъй като интернет технологиите се развиват бързо, все повече и повече данни могат да бъдат кодирани в нестандартни форми, които са трудни за извличане - аудио- или видеофайлове, засилено използване на интерактивно или потребителско специфично съдържание. Ето защо, някои от уебсайтовете - по-специално на големите предприятия или тези, работещи в творческите индустрии, потенциално могат да станат трудно достъпни и невъзможни за извличане на информация. Поради тази причина е възможно да се наложи повторно извършване на „web-scraping“ на данни от висококвалифицирани специалисти във или извън НСИ, които ще изискват и по-голям финансов ресурс.

Практическата реализация на проекта поставя редица въпроси пред специалистите от НСИ, свързани с необходимостта от конструиране на сложните данни, извлечени в голям мащаб, преди по-нататъшен анализ, като прилагането на машинно обучение за валидиране, свързване и интегриране на данни. Не по-малко важен е проблемът с извършването на „web-scraping“ от онлайн машини в мрежата и прехвърлянето на извлечените данни от офлайн мрежи в онлайн хранилища. НСИ се нуждае от сходни системи за съхранение на данни, които могат да



управляват целия жизнен цикъл на данните - съхранение, проектиране и поддържане на големи бази.

В заключение може да се каже, че разглеждайки Big Data с тяхната реална стойност и значимост и връзките им с други явления, смисълът на тяхното използване като един от възможните източници в официалната статистика изглежда по-скоро логичен, отколкото необичаен. Националните статистически институти трябва да имат фундаментални знания и да разширяват опита си по отношение на използването на Big Data в ежедневната статистическа практика и извън нея. Прилагането на принципа „количество над качество“, възприет от потребителите на Big Data, не трябва да се пренебрегва. Дори когато източниците на Big Data не се използват за получаване на нови статистически продукти, те биха могли да се разглеждат като ефективно средство за намаляване на натовареността на респондентите, при условие, че методологичните предизвикателства могат да бъдат разрешени. Използването на Big Data за съставяне на ранни показатели за важни статистики като например данни за цените или бизнес цикъла е достатъчно сериозна опция. Прилагането на Big Data за краткосрочни прогнози също не е за пренебрегване.

В крайна сметка полетата на теорията и практиката са обединени от една цел - получаване на реални и навременни изводи от публично и лесно достъпни данни. Развитието на обществото и информационните технологии разкриват нови потребителски очаквания, обекти и феномени на интерес, за които официалната статистика не е в състояние да предостави данни. „Големите данни“ са съществена част от това развитие. В този смисъл може да се добави, че чрез използване на принципите на нанотехнологиите Big Data ще могат да се метрират и това ще даде облика на 21-ви век.

ИЗПОЛЗВАНА ЛИТЕРАТУРА:

Ангелова, П. (2013). Статистика. Свищов, СА „Д. А. Ценов“.

Богданов, Б., Г. Статева (2017). Бъдещето на изследванията и изследванията на бъдещето: възможни приложения на големите данни при производството на статистическа информация. Статистика, № 1.

Европа 2020 (2010). Стратегия за интелигентен, устойчив и приобщаващ растеж. Европейска комисия, Брюксел, 3 март, COM(2010) 2020 окончателен. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:2020:FIN:BG:PDF>.

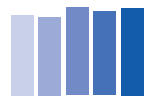
Alpaydm, E. (2004). Introduction to Machine Learning (Adaptive Computation and Machine Learning). MIT Press, ISBN 0-262-01211-1; Witten, I., Frank, E., Hall, M. (2011). Data Mining: Practical machine learning tools and techniques. 3rd ed. Morgan Kaufmann. ISBN-13: 978-0123748560.

Boeing, G., P. Waddell (2016). New Insights into Rental Housing Markets across the United States: Web Scraping and Analyzing Craigslist Rental Listings. Journal of Planning Education and Research, 23 august. <http://journals.sagepub.com/doi/pdf/10.1177/0739456X16664789>; Vargiu, E., Urru, M. (2013). Exploiting web scraping in a collaborative filtering- based approach to web advertising. Artificial Intelligence Research, vol. 2, №1. <http://www.sciedu.ca/journal/index.php/air/article/view/1390>.

Buelens, B., P. Daas, J. Burger, M. Puts, J. Brakel (2014). Selectivity of Big data. Discussion Paper, №11, Statistics Netherlands. <file:///D:/2014-11-x10-pub.pdf>.

Daas, P., M. Puts, B. Buelens, P. van den Hurk (2015). Big Data as a Source for Official Statistics. Journal of Official Statistics, Vol. 31, No. 2, pp. 249-262. <http://dx.doi.org/10.1515/JOS-2015-0016>.

Eurostat (2015). Experiments for using big data in official statistics. 2015 UNECE Project.



ПРИЛОЖНИ АСПЕКТИ НА „ГОЛЕМИТЕ ДАННИ“ В ОФИЦИАЛНАТА СТАТИСТИКА

Галя Статева*

РЕЗЮМЕ Настоящата статия е предназначена да запознае читателя с една възможност за практическо приложение на „големите данни“, която е реализирана в рамките на проведено от екип на НСИ емпирично изследване на тема „Извличане на информация от интернет за характеристики на предприятията (web-scraping)“.

Във Въведението подробно е обяснена *основната цел* на проведеното емпирично изследване, която е насочена към проучване на възможностите за прилагането на техниките „web-scraping“ и „text mining“ и оценяване на ефекта от използването им в процеса на събиране на данни и подобряване на качеството на информацията за предприятията от статистическия бизнес регистър на НСИ чрез достъп до техните уебсайтове.

В глава I подробно е представена технологичната среда за осъществяване на емпиричното изследване, включително общата референтна логическа архитектура за прилагане на техниките на „web-scraping“. Направена е обстойна характеристика на техниката „web-scraping“ и са описани случаите, в които се използват различните видове - специфичен и генеричен „web-scraping“.

Изложението в глава II е посветено на практическата реализация на четирите пилотни „сценария“. Провеждането и анализът на резултатите от отделните „use-case“ са представени в аналогична последователност - цел, ресурсна и технологична осигуреност, постигнати резултати, правни ограничения.

* Д-р, държавен експерт в отдел „Обща методология и анализ на статистическите изследвания“, дирекция „Методологично-учебен център“, НСИ; e-mail: GStateva@nsi.bg.

ПРИКЛАДНЫЕ АСПЕКТЫ БОЛЬШИХ ДАННЫХ (BIG DATA) В ОФИЦИАЛЬНОЙ СТАТИСТИКЕ

*Галя Статева**

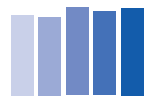
РЕЗЮМЕ В настоящей статье представлена возможность практического применения Больших данных (Big Data), реализованной в рамках эмпирического исследования „Извлечение данных из интернета для получения информации о предприятиях (web-scraping)“, проведенного Национальным статистическим институтом.

Во введении подробно описана основная цель эмпирического исследования, которая направлена на изучение возможностей применения методов „web-scraping“ и „text mining“ и оценки результатов, полученных от их использования в процессе сбора данных и улучшения качества информации о предприятиях в статистическом бизнес-регистре НСИ, через доступ к их веб-сайтам

В первой главе представлена технологическая среда для реализации эмпирических исследований, включительно общая референтная логическая архитектура для применения метода „web-scraping“. Дается подробная характеристика метода „web-scraping“ и описаны случаи, в которых используются различные виды – специфический и генерический „web-scraping“.

Вторая глава посвящена практической реализации четырех пилотных „сценариев“. Разработка и анализ результатов отдельных „use-case“ представлены в следующей последовательности: цель, ресурсное и технологическое обеспечение, достигнутые результаты, юридические ограничения.

* Д-р, государственный эксперт в отделе „Общая методология и анализ статистических исследований“, дирекция „Учебно-методологический центр“, НСИ; e-mail: GStateva@nsi.bg.



APPLICABLE BIG DATA ASPECTS IN OFFICIAL STATISTICS

*Galya Stateva**

SUMMARY This article is intended to acquaint the reader with an opportunity for practical application of ‘BIG DATA’, which was realized in the framework of an empirical research carried out by an NSI team on the topic of ‘Extracting information from the Internet for business features of enterprises (web-scraping)’.

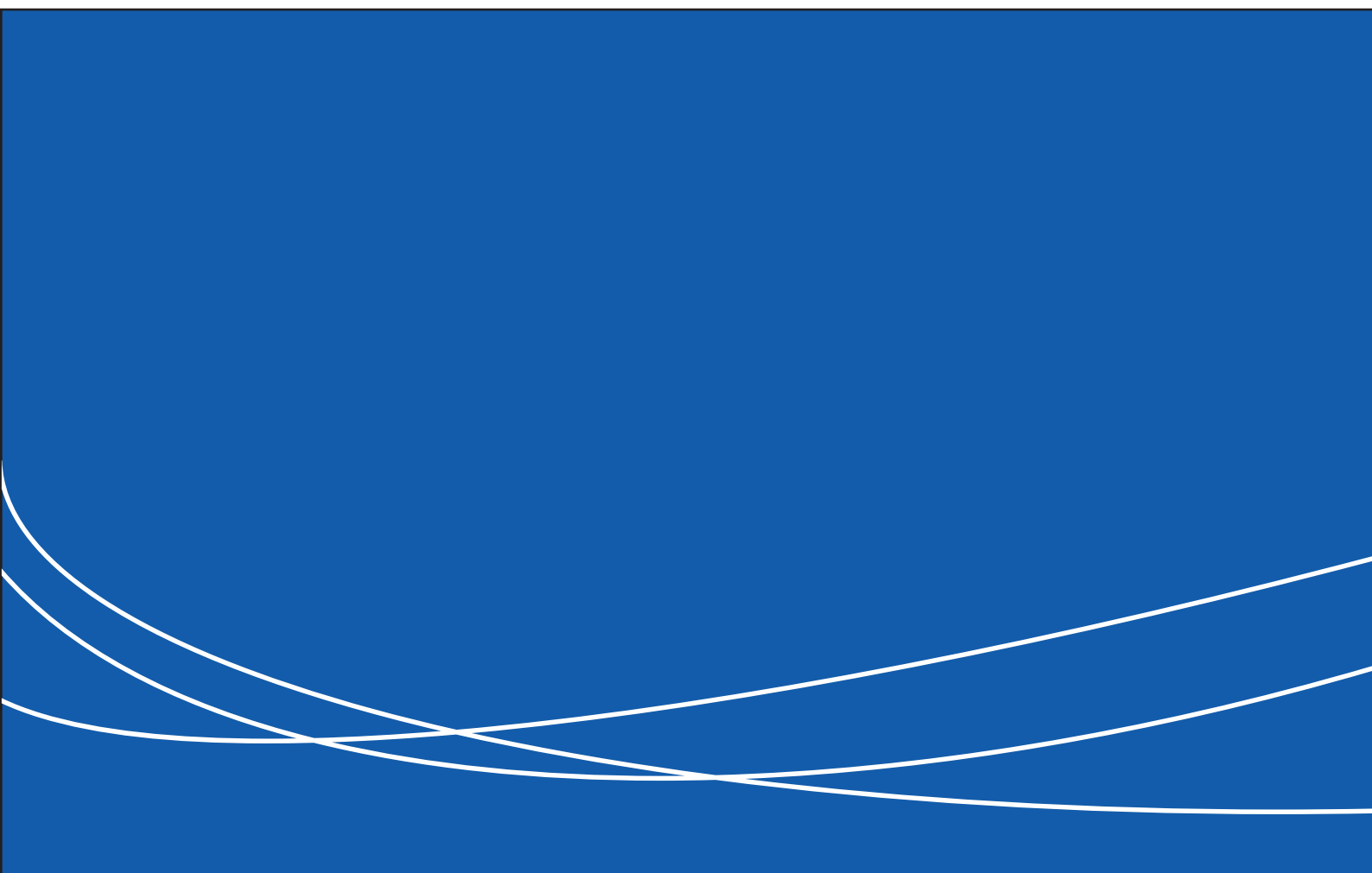
The Introduction of the article explains in detail *the main purpose* of the conducted empirical study, which is aimed at exploring the possibilities of applying ‘web-scraping’ and ‘text mining’ techniques and evaluating the effect of their use in the process of data collection and quality improvement of enterprise information from the NSI Statistical Business Register through access to their websites.

Chapter I presents in detail the technological environment for the implementation of the empirical research, including a common reference logical architecture for the application of ‘web-scraping’ techniques. A detailed characteristic of the ‘web-scraping’ technique is presented and the cases in which various types of specific and generic ‘web-scraping’ are used is made.

The presentation in Chapter II is devoted to the practical realization of the four pilot ‘scenarios’. The conduction and analysis of the results of the various ‘use-cases’ is presented in an analogous sequence - objective, resource and technological provision, results achieved, and legal constraints.

* Dr, State Expert in the General Methodology and Analysis of Statistical Surveys Department, Methodological and Training Center Directorate, NSI; e-mail: GStateva@nsi.bg.

**СТАТИСТИЧЕСКИ ИЗСЛЕДВАНИЯ И
АНАЛИЗИ**



**АКАДЕМИЧНИ НАУЧНИ ПОЗИЦИИ ЗА СТАТИСТИЧЕСКАТА НАУКА И ЗА
ПРОМЕНЕТЕ В ПРОФЕСИЯТА НА СТАТИСТИЦИТЕ В
СЪВРЕМЕННИЯ СВЯТ**

*Димитър Радилов**



Въведение

Предметът на статистиката като наука в съвременния свят все още е обект на оживени научни дискусии на различни научни конференции, организирани от Международния статистически институт (МСИ), академични творчески институции и неправителствени организации. Схващането, че статистиката е клон на математиката, все още не е изживяно. Не е случайно, че на V сесия на МСИ в Хелзинки (Финландия) през 1999 г. се обсъждат решаващите изходи за дългосрочно оцеляване на статистиката, както и на професията в близките двадесет години.

Научната цел на статията е разглеждането на промените в академичните научни позиции за статистическата наука и необходимите промени в професията на статистиците в съвременния свят.

1. Предизвикателствата на съвременния свят за промени в академичните научни позиции за статистическата наука

В съвременния свят централно място заемат информацията, знанията и комуникациите. Промените в него могат да се представят според мен по следния начин:

* Проф. д-р ик. н. в катедра „Статистика и приложна математика“, Икономически университет - Варна;
e-mail: d.radilov@gmail.com.

Първо, съвременният свят се развива динамично и шоково като сложна и многопластова социална структура.

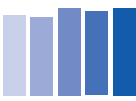
Второ, създаването на нови информационни технологии (компютър, интернет, мобилен телефон, сателитни чинии, облачни технологии и др.) и подобряването на информационните инфраструктури водят до съществена промяна в живота, поведението и ценностната система на човешкия фактор. Това са предизвикателствата за промени.

Трето, глобализацията като основен феномен завладява съвременния свят. За него няма граници. Възникват нови направления на науката „глобалистика“: 1) футурологична глобалистика (наука за запазване на земната цивилизация); 2) геоекология; 3) геоекономика; 4) социално-демографска глобалистика и 5) геополитика. Този процес е необратим. Подчинен е на логиката на новите информационни и комуникационни технологии. Това са предизвикателства към глобализацията.

Четвърто, науката и образованието са едни от основните ценности на съвременния свят. Непрекъснатите промени изискват адаптивни хора, които бързо навлизат в най-новото и трябва да проявяват креативност, а не само изпълнителност. Това са предизвикателства за промени в професиите и длъжностите на трудовия пазар. Производителността се определя като начин на мислене, вяра в науката и образованието за превръщане на университетските знания в иновации.

Пето, в съвременния свят най-трудно е човек да намери своето място. Новите информационни технологии променят много бързо знанията, които се търсят. На трудовия пазар най-много се ценят интелектуалните качества, натрупаните знания и умственият труд за превръщането им в иновации. Университетската диплома е само етап от ученето през целия живот. Това е предизвикателство към всяка професия, в т.ч. и статистическата.

Академичните научни позиции за статистиката като наука в международен план са свързани с UNESCO, Международния статистически институт, Кралското статистическо общество във Великобритания и Американската статистическа асоциация. Счита се, че те в една или друга степен са повлияни от развитието на предизвикателствата на съвременния свят в края на XX и началото на XXI век. В новото издание на Оксфордския речник на статистическите термини (съвместно с МСИ) „статистиката е наука за събиране, анализиране и тълкуване на данни, обединени в агрегати“ (The OXFORD Dictionary of STATISTICAL TERMS, 2003, p. 388). В



международната стандартна класификация на образованието на ЮНЕСКО статистиката се дефинира като „наука за събиране и групиране (collect), описване, подреждане и анализиране на данни“ (Eurostat, Fields of Education and Training, 1999, pp. 3, 18). Научната позиция на проф. Д. Линдлей в монография, публикувана в списанието на Кралското статистическо дружество във Великобритания, е, че „статистиката е „наука на неопределеността“. Тя „се опира на три принципа: първо, че неопределеността трябва да се измерва с вероятности; второ, че резултатите от измерванията трябва да се оценяват чрез полезности, и трето, че оптималното решение обединява резултати, получени от измерване на неопределеността с вероятности с тяхната полезност“ (Lindley, Denis V., 2000, pp. 293, 318).

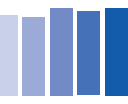
Една по-задълбочена и критична оценка на тези три определения на „статистическата наука“ като понятие изисква да отбележим най-малко три методологически особености. **Първата** от тях е свързана с признаване на информационното съдържание на статистическата наука в определенията на Оксфордския речник на статистическите термини, МСИ и ЮНЕСКО. Това е крачка напред. **Второ**, всяко от цитираните определения само по себе си отразява съдържанието на понятието „статистическа наука“ само по отношение на отделни негови компоненти (концептуални, математически, информационни и т.н., което произтича от възприети частно-научни подходи). Това е крачка встрани. **Трето**, наблюдава се процес на методологична фрагментарност, която намира израз в различни абсолютизации на отделни аспекти на статистическата наука, например определението на Д. Линдлей, като не се отчита необходимостта от цялостност и комплексност, съобразена със съвременното развитие на науката като социална система. Това е крачка назад. Втората и третата методологична особеност не стимулират едно развиващо се по теоретично съдържание определение на статистическата наука. И нещо повече, при преглед на литературата, посветена на теоретико-методологичното съдържание на статистическата наука в исторически аспект, се вижда, че статистическата наука подобно и на други науки, независимо дали са хуманитарни, или природни, все още няма общоприет (и всепризнат) теоретичен фундамент и прецизно определен предмет на изследване със структуриране на включените в него понятия, теоретични, методологични, аналитични и познавателни инструменти.

При едно познавателно тълкуване на съдържанието на понятието „статистическа наука“ винаги съществуват непредвидими рискове за неговото оспорване и превръщане в трудно решим проблем. Подобно негласно оспорване на статута на статистиката като

наука се отразява върху имиджа и традиционните позиции на статистическата наука сред информационните науки. „Доминиращите перспективи на тези загуби са свързани с информационния пазар. Доскоро статистиката според Ж. Фридман беше една от малкото науки, която се занимава приоритетно с анализ на данни. Днес в условията на експоненциално развитие на информацията и информационните технологии, свързани с нея, има много други науки, които са приоритетно ориентирани към анализа на данни. Те се съревновават успешно за клиенти, перспективни изследователски обекти, специализации и магистърски програми, студенти и ценни с висок социален статус работни места и в много случаи печелят“ (Friedman, J. H., 2001, pp. 5 - 10).

На фона на цитираните определения на статистическата наука и нейното място сред информационните науки според моето виждане **важно място заемат няколко предпоставки. Първата** от тях е, че в хода на историческата еволюция, знанията на статистическата наука са били използвани като фундаментална база за достоверна информация за управлението. **Втората** е, че развитието на статистиката като наука се съпътства от разнообразни теории (икономически, политически, математически, философски и т.н.), които пряко или имплицитно оказват влияние върху определянето на нейното съдържание по същество. **Третата** е, че липсата на ясна дефиниция на понятието „статистическа наука“, което пречи на развитието на системата на нейните знания през XXI век и още повече при определяне на нейното място в съвременния свят. Така логически стигаме до необходимостта за определяне на обхвата и структурирането на знанията на статистическата наука, макар и в най-общ теоретичен план. Обхватът и структурирането на статистическата наука, оценени от „птичи поглед“, са колкото важни, толкова и почти неразработени в специализираната статистическа литература.

Индийският статистик проф. Махаланобис в едно свое изследване за статистическото обучение в социалните науки през 1957 г. предлага двете направления на статистиката, едното от които се занимава със събиране и анализ на информация за правителството и за цели на планирането, а другото - с моделиране и правене на изводи от разнородна информация, да се слоят (Mahalanobis, P. C., 1957). Във Франция и Германия тези съставни части са налице, но подобно сливане не се е осъществило. В англоговорещите страни се отдава по-голямо значение на приложните аспекти и предметът на статистиката се разглежда повече като самостоятелна унифицирана учебна дисциплина. Тези различия в развитието на статистиката не са случайни, а имат дълбоки корени в институционалното и образователното развитие на отделните



държави. Тази визия на Махаланобис за сливане на официалната статистика с академичната статистическа наука позволява на Индия още през 1960 г., макар слаборазвита и обхваната от кризи страна, да изнася първокласни статистики за водещи университети в САЩ. За проф. Махаланобис статистиката не е цел, а средство към целта, най-мощното достъпно оръжие за постигане на силно желани социални, национални и хуманитарни цели (Vere Jones, D., 2002, p. 40).

Според проф. Никълъс Фишер (Fisher, Nicolas, 2001, pp. 11 - 15) развитието на статистиката в перспектива изисква отговор най-малко на следните въпроси:

а) как можем да направим така, че статистиката да се счита не само като уместна, а като жизненоважна в публичното пространство на дискусиите и обществения живот, като се промени положително отношението към нея и статистиците?

б) как можем да привлечем най-доброто и перспективното в съдържанието на статистиката? Статистиците трябва да запазят и развият своя професионален образ, различен от идентифицирането им като лица, които само ползват математиката. Преподаването на статистика трябва да научи статистиците да комуникират в публичното пространство с нестатистиците чрез разширяване на връзките им и въвеждане на екипност в тяхната дейност.

2. Категориите „съвкупност“ и „множество“ в статистическата наука като две страни на единното битие на масовите явления, процеси и действия

В международната стандартна класификация на образованието статистиката, информатиката и математиката са поставени в една и съща група. Общото между тях е, че имат отношение към създаването, анализа и използването на статистическата информация. В случая това, което разграничава статистическата наука от тези две науки е специфичното понятие „съвкупност“, което липсва в тях.

Съвкупността е основно понятие в статистическата наука. Тя има материално-веществено съдържание за реалните масови явления и процеси. Означава се със символа „G“.

Понятието „масово явление“ е свързано с голям брой единици, които са обхванати в статистическа съвкупност по даден признак или признаци в конкретни времеви и пространствени граници. С тях се свързват масовите явления, процеси и действия. Те не са нови дадености, качествено или количествено различни от единиците, които ги съставят и които се регистрират при масовото наблюдение във връзка със статистическото изучаване. В случая изучаваните масови явления са само

продукт на една логическа абстракция при обхващането им в съвкупности. Тя позволява да бъдат концептуално измерени и анализирани чрез подходящи обобщаващи характеристики за получаване на нови знания.

Закономерността изразява наличие на причинно-следствена връзка. Статистическите закономерности се изследват само в масови явления и процеси. Свойствата на статистическите закономерности се разкриват при обобщаване на данни за достатъчно голям брой единици. Те винаги се свързват със Закона за големите числа. Със статистическия метод се разкрива действието на една или друга статистическа закономерност в конкретните условия на дадено място или време с особености, които трудно могат да се предвидят, ако не са налице числени данни. Той има две функции - описателна и аналитична, които липсват във формалната логика.

В исторически план понятието „съвкупност“ е въведено от Кристов Зигварт (1830 - 1904) в неговата „Логика“ (Г. Данаилов, 1932, с. 135 - 138) под названието „колективно понятие“. По-късно то се развива от Г. Кнап (1868) и Е. Рениш (1883), разбира се, с някои особености. Под „колективно понятие“ К. Зигварт разбира едно цяло, съставено от предмети или индивиди, които са множество дискретни части, мислими като отделни единици (пак там, 1932, с. 135).

В английския език за понятието „съвкупност“ през 1915 г. А. Боули въвежда термина „популация“ (population). Той счита, че „популациите включват единици, които притежават общ признак, но се различават количествено по своята величина“. В най-новия Оксфордски речник на статистическите термини от 2003 г. понятието „популация“ (съвкупност) съдържателно се свързва с термините „вселена“ (universe) и „агрегат“ (aggregate). В статистически смисъл понятието „популация“ (съвкупност) се прилага за всяко крайно или безкрайно натрупване (струпване) на хора. То е заменило по-стария термин „вселена“, който е извлечен от „вселената на разсъжденията“ в логиката. На практика негов синоним е „агрегат“ и не е задължително да се отнася за натрупване (струпване) на живи организми.

М. Кендал и А. Стюарт в своята монография „Теория на разпределенията“ определят понятието „групи“ или „семейства“ като фундаментално в статистическата теория (Кендалл, М., А. Стюарт, 1966, с. 39, 287 - 288). Те дават едно много широко определение на понятието „природен феномен“, включващо всички възможни обекти на външния свят, които са свързани или не са свързани с човека.

Новото, което проф. Г. Данаилов прибавя към знанията за статистическата съвкупност, е, че тя е изкуствено създадено колективно понятие, което има за обект



масови явления, за които всяко едно число, получено чрез преброяване, има реална ценност по време и място (Данаилов, Г., 1932, с. 622 - 623).

Определение на понятието „съвкупност“ дава и проф. Венец Цонев. Бих искал да отбележа, че етимологията, смисловото развитие и понятийната изразителност на думата „съвкупност“ са изследвани задълбочено от проф. В. Цонев (Цонев, В. и С. Бадева-Цонева, 1983, Статистика, кн. 1, с. 13 - 23). Той „разглежда съвкупността като цяло, съставено от физически отделени една от друга единици, имащи материално-веществено битие, например човешки индивиди, отличаващи се с пространствена мобилност“. Проф. В. Цонев се спира и на осезаемостта на единиците на съвкупността. С тези свои научни позиции той обобщава и допълва определенията на Г. Кнап (1868), Е. Рениш (1883) и на Г. Данаилов (1932).

Възприетият от мен подход е доразвитие на обобщенията на проф. В. Цонев. Той се основава на моята теза за структуриране на знанията за съдържанието на статистическата съвкупност като концептуални, аналитични и тълкувателни. Трите вида знания за статистическата съвкупност при измерване на масови събития имат аналогично значение за статистическата наука както възприетия, по думите на акад. Б. Петканчин, аксиоматичен подход при използването на понятието „множество“ в математиката. Именно чрез разработването на различни концепции за опосредствано измерване на масови явления се реализира идеята за създаване на съвкупностна информация с използване на разнообразни количествени инструменти и тяхното тълкуване за създаване на нови знания и иновации.

Статистическата съвкупност според моята позиция е цяло от много самостоятелни (дискретни) материално-веществени елементи (предмети, индивиди, организми, институционални единици и др.), обединени по общ признак, които са свързани със случайни, причинни или целесъобразни връзки при изучаване на крайни масови явления във времето и пространството.

Понятието „множество“ е първично в математиката. Под множество разбираме което и да е обединение в едно цяло M на определени напълно различими обекти (наричани „елементи“ на M) на нашето възприятие или мисли. Знанията за множеството се използват чрез аксиоматичен подход в математиката (Kleene, S., New York, 1952).

Първата система аксиоми на теорията на множествата (Александрова, Н. В., 1978, с. 79 - 80) е разработена през 1904 - 1908 г. от немския математик Е. Цермело (1871 - 1953). В нея, „грубо казано, имаме първични понятия: обекти, съвкупност, \in

(елемент на съвкупност), свойства и 9 аксиоми. Те са изрично отбелязани като аксиоми. Няма да се спираме на въпроса дали може да се намали броят на първичните понятия и аксиомите. От системата на Цермело не сме дали например аксиомата за безкрайността, за която ще стане дума по-късно. Ролята на аксиомите е главно: 1) да се постулира съществуването на някои прости съвкупности, образувани по дадени обекти (с. 1 - 3); 2) да се дадат начини за образуване на съвкупности [по дадени признаци] (с. 4, 9, 24, 34). Така се ограничава преголямата свобода в образуването на съвкупности, благодарение на която именно се стига до логически трудности“ (Петканчин, Б., 1968, с. 5 - 6, 13). Приносите на Е. Цермело са насочени към постигане на изчерпателност, строгост и систематичност в аксиоматичния подход, приложен в теорията на множествата. По този начин се открива изход от кризата и насока за по-нататъшно развитие на теорията. Започва триумфалното шествие на теорията на множествата във всички области на математиката.

В теорията на вероятностите множеството при аксиоматичното му използване е един елемент (Ω) от вероятностното пространство (Димитров, Д., 2004, с. 16), на основата на което се основават класическата вероятност, случайните величини, функциите на разпределение на вероятностите, на плътността им, Законът за големите числа, теоретичните разпределения и т.н.

В математическата статистика „множеството“ се представя с генерална съвкупност и извадково пространство. Те се използват за решаване на задачи за оценка на вероятности, за проверка на хипотези, за емпирични функции на разпределения и т.н. (Димитров, Б. и Н. Янев, 1998, с. 167 - 173).

Категориите „съвкупност“ и „множество“ са взаимосвързани. Те дават различни представи при измерване на масови явления в единното битие на света. „То е реална действителност. Има вътрешна логика на развитие. Обхваща природата като цялостен универсум, който е бил, съществува и ще бъде. Човекът и обществото, които някога са възникнали, съществуват и ще съществуват още дълго време. Човекът е носител на съзнание. То е осъзнато битие“ (Марчев, Ф., 2003, с. 272 - 273). Научното определение на битието „според принципите на философията“ включва две съставки. Те са природата и обществото. **„Природата като част от битието е цялата безкрайна по пространство и време качествено многообразна реалност.** Тя обхваща най-общо механичната, физичната, химическата и биологическата природа на хората, както и природните свойства и закономерности на преработените в процеса на производството природни предмети, с които хората си служат в производството. **Обществото като**

част от битието е реалният процес на техния живот. То включва: 1) цялата обективно съществуваща човешка практика, т.е. производството, разпределението, потреблението и натрупването на блага; 2) управленската дейност, социално-икономическите и политическите движения и 3) цялата институционална система от предприятия и организации на социалния живот, т.е. семейство, домакинство, гражданско общество и държава“ (Марчев, Ф., цит. кн., с. 280). При това природата може да съществува без общество, например континентът Антарктика, който е част от южната полярна област от Земята (Българска енциклопедия, БАН, 2002, с. 56). Обществото не може да съществува без природата. Тя е негова среда за възникване, съществуване и развитие. Обществото е част от нея.

3. Статистическата наука като социална система от знания

Идеята за развитието на статистиката като социална система от знания се поставя за първи път от Ламберт Адолф Кетле през 1836 г. в изследването „Социална физика или опит за изследване развитието на човешките способности” (1836). Той поставя началото на философските изследвания за разкриване на законите, които управляват света (Плошко, Б. Г. и И. И. Елисева, 1990, с. 32 - 40).

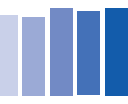
Проф. Георг фон Майер написва фундаменталния научен труд „Статистика и учение за обществото” (1895 - 1909) (Майер, Г., 1899, с. 9, 24 - 26). В нея защитава тезата, че статистиката е наука за социалните маси (хора, действия и резултати от човешки действия). Според големия теоретик логик Франц Джиджек развитието на статистиката е синтезирана социална наука (Стефанов, Ив., 1940, с. 46 - 47). Уилям Петти е първият, който изчислява националното богатство и националния доход на Англия и Уелс (Петти, У., 1940, с. 17). Известният статистик М. Ханзен (1910 - 1990) разработва методология за репрезентативни изучавания на паралелни проучвания при преброяване на заетите лица и безработните през 1937 г. (Чолаков, Н., 2007, с. 437 - 440). Теоретичните и практико-приложните приноси на проф. Оскар Андерсон (1887 - 1960) са обобщени в монографията му „За репрезентативния метод и неговото приложение при разработката на материала за преброяване на земеделските стопанства към 31.XII.1926 г.”. Той е пионер в използването на репрезентативния метод в земеделската статистика (Андерсон, О., 1929, с. 109 - 152).

Според проф. Залм пропуските и недостатъците в статистическата информация, включително проблемите по внедряването на световните стандарти (Система за националните сметки) на една широка основа, се сблъскват с нерешени концептуални въпроси. След това той акцентира на напрежението, което произлиза от факта, че

„светът се променя”, „теорията се променя” и „средствата се променят”. В проекта „Изследването на бъдещето” акцентите според проф. Залм по отношение на основните бъдещи насоки на статистиката са следните: а) нуждата да се осигурят статистически данни, които са необходими за един по-широк обзор на обществените процеси с цел по-добре да бъде разбран напредъкът в икономиката; б) нуждата от по-подробна информация с по-гъвкави възможности за сумиране; в) нуждата да се допълнят количествените данни с качествени за структурата на пазара, за институциите, за риска и за несигурността; г) нуждата да се допълни количествената информация за статичните условия, които представляват баланс с данни по отношение на динамичните характеристики на вечно изменящите се процеси; д) нуждата от подсилване и разширяване на сравнимостта и последователността на статистическите данни в световен мащаб (Kenessey, Zoltan, 1994, pp. 1 - 14).

В контекста на съвременния свят статистическата наука ще се развива като информационна социална система от знания най-малко в следните направления:

- **първо**, статистическата наука трябва да развива концептуални знания за измерване на масови събития в реалния свят с числена информация, която да има потребители, неоспорвани от другите науки. В този смисъл едната ѝ разновидност са концептуалните знания;
- **второ**, за да изпълни своята методологична функция за превръщане на създадената числена информация в нови знания, статистическата наука трябва да използва разнообразни количествени инструменти за решаване на познавателните си цели и задачи, главно от математиката;
- **трето**, за да изпълни своята социална познавателна функция, статистическата наука трябва да насочи развитието на знанията си към тълкувателни подходи за грамотност и съвкупностно мислене в публичното пространство при разпространение и използване на създадената информация и нови знания от нея като иновации;
- **четвърто**, необходимо е концептуалните знания на статистическата наука да се насочат към измерване на масови явления в икономиката, финансите, социалната сфера и новата икономика, доколкото те са приоритети в развитието на съвременния свят чрез създаване на статистически информационни системи;
- **пето**, знанията на статистическата наука трябва да се използват за засилване на мястото ѝ в развитието на статистическата дейност чрез промени в съдържанието на



статистическата професия и образование и при продължаващо обучение през целия живот на нейните субекти.

Предлагам едно ново определение за обекта и предмета на статистиката като система от научни знания. Обект на статистическата наука са масовите явления на човешкото общество в неговата природна среда. Предмет на статистиката като наука е изучаване на закономерностите в масовите явления чрез система от концептуални, аналитични и тълкувателни знания, основани на статистическата съвкупност. Лансирам виждането, че в бъдеще част от статистическата наука ще се развива и като информационна.

Приоритетите на статистическата наука през XXI век в новите реалности на съвременния свят са насочени към усъвършенстване и развитие на теорията ѝ. Тезата ми е, че структурирането на теорията на статистиката като наука включва три типа знания: концептуални, аналитични и тълкувателни, които се развиват върху логиката на познавателния процес за съвкупности. Те са в унисон с разглежданите приоритети. Разлагането на знанията на статистическата наука на три типа според мен в случая се използва като форма на научна абстракция за прецизиране на понятията „теория“ и „методология“. В реалната познавателна дейност не могат да бъдат пренебрегнати връзките между концептуалните, аналитичните и тълкувателните знания. Въвеждането им се основава на мотива, че в някаква степен те могат поотделно да са структуроопределящи в „теорията“ и „методологията“. Не ми е известно друго подобно структуриране на знанията на статистическата наука.

4. Статистическата професия в съвременния свят

Мястото на статистическата професия в съвременния свят се разглежда за първи път на Общата асамблея на Международния статистически институт през 1987 г. в Токио. На нея Джон Гоуър, ръководител на отдел по „биостатистика“ в Ротънстедската експериментална станция в Обединеното кралство, и Питър Дигел, ръководител на отдела по „Математика и статистика“ (CSIRO), правят следното изказване:

„Има признаци, че в условията на финансов стрес ръководителите на изследователски проекти често възприемат статистическите дейности като област, в която могат да се направят икономии. Това може да се дължи на ред причини:

- Незнание за това, което правят статистиците, особено що се отнася до разходната ефективност на доброто проектиране и безпристрастен критичен анализ;
- Погрешно схващане, че комплексните компютърни програми елиминират нуждата от професионален статистически съвет;

- Поради относително малкия брой статистики в сравнение с другите професионални групи гласовете на статистиците често биват пренебрегвани при вземане на важни политически решения.

Положението е особено сериозно в промишлеността, в изследователските институти и в университетите. Ние искаме МСИ да обърне внимание на това положение и да прецени дали е в състояние да сформира работна група, която да формулира практически стъпки за разрешаване на проблема. Трябва да се разработят методи, които да гарантират, че висшите ръководители и овластените да вземат решения разбират по-добре предимството на науката и на възможностите да се наемат творчески настроени статистики в промишлеността в реално действащи професионални групи“ (Moriguti, S. et al., 1992, p. 227).

Сформирана е работна група с ръководител проф. Сигети Моригути, която да проследи проблемите, засягащи професията, поради неправилното схващане за полезната роля и положителния принос на статистиците в съвременното общество. Правилното разбиране на ролята на статистиците изисква комплексен системен анализ на всички области, посочени във фиг. 1.

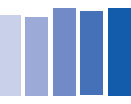
В разработения доклад са формулирани 7 препоръки. Те са:

Препоръка 1. МСИ трябва да действа като център на разбирането, че статистическата система е основна част от „Информационната инфраструктура“ на света.

„Статистиката изпълнява съществена роля в много области на обществения живот. Освен това в силно взаимозависимия съвременен свят познаването му като цяло е съществено за вземането на правилни решения не само за микрополитическата сцена, но дори и на относителни макроравнища. В този смисъл статистиката представлява същностна част от информационната инфраструктура. Това означава, че статистическата информация трябва да се развива като система, която да обхване с достатъчна прецизност целия свят - Севера и Юга, Изтока и Запада“ (Moriguti, Sigeiti et al., 1992, p. 229).

Препоръка 2. МСИ трябва да стимулира съответните действия на държавите в зависимост от степента на развитие на статистическите им системи.

„Добре действащата статистическа система се създава в продължение на много години. Тя също така трябва често да се оценява, подобрява или модифицира, за да отговаря на динамичните нужди и ситуации в обществото.



Ето защо е препоръчително да се поддържа дългосрочна статистическа програма във всяка страна или регион (включващ подобни или близки в географско отношение страни). Дори в развитите държави една дългосрочна стратегия за развитие на официалната статистика служи като мощно средство за координиране на статистическите дейности на различни агенции (министерства, отдели и т.н.) и спомага за осигуряването на необходимите бюджетни средства.

Голяма е грижата и за образованието и обучението по статистика, тъй като ефектът от тях е осезаем дълго след като са взети или не правилните действия. Продължаващото обучение, квалификация и преквалификация на заетите вече в пазара на труда са също така важни, особено в бизнеса.

Тъй като статистическата система е съществена част от информационната структура на обществото, нейното развитие и поддържане е отговорност на обществото, най-вече на национално ниво. Ако правителството изпитва недостиг на ресурси - човешки, финансови или други, то трябва да се осигури международна помощ по един разумен начин. Едни общи ръководни принципи, съобразени с нивото на развитие, биха подпомогнали тези усилия, макар специфичните задачи да се различават в отделните страни” (Moriguti, S. et al., 1992, p. 230).

Препоръка 4. МСИ трябва да стимулира сътрудничеството между статистици и компютърни специалисти.

„Връзката с компютрите или компютърните специалисти е изключително важна. В днешно време един статистик, неспособен да използва компютри и статистически софтуер, е в крайно неблагоприятна позиция. Естествено, за да се превърне в желан партньор, статистикът трябва да притежава редица други качества. Познаването на компютрите днес е задължително изискване към всички статистици” (Moriguti, S. et al., 1992, p. 231).

Препоръка 5. МСИ трябва да засили чрез ресурси връзките на статистиците с масмедията и с обществеността.

„Ключовите фигури, вземащи решения в правителството, в индустрията, в бизнеса и в науката, са стратегически важни за засилване ролята на статистиката и на статистиците.

Политиците и висшите държавни служители, които вземат най-важните решения в управлението, се влияят директно от разбирането и подкрепата или от неразбирането на обществеността. Затова правилното тълкуване на статистическите данни и

обоснованото мнение за тях са съществени в достигането на добри политически решения.

Медиите, т.е. вестниците, телевизията и др., представляват мощно средство за образование на обществото по отношение на статистическите данни. Трябва да се положат адекватни усилия за това медиите да тълкуват правилно и разбираемо статистическите проблеми.

В сферата на бизнеса е важно заетите и работодателите да разберат ролята на статистическото мислене и на статистическите методи. Това може да стане с помощта на специални квалификационни курсове и обучение на работното място.

МСИ трябва да разширява своите двугодишни сесии с броя на срещите по теми, които са предизвикали широк интерес, и да играе положителна роля в предаването посредством медиите и пресинформациите на съществени решения по тези въпроси” (Moriguti, S. et al., 1992, pp. 232 - 233).

Препоръка 6. МСИ трябва да засили публикационната активност на статистиците по отношение на мислене, методи и изследвания.

„Макар най-ефективните инициативи да се провеждат на национално ниво, МСИ може да предостави ценна информация за тях. Чрез своите известия Институтът оповестява съществуващи национални инициативи и засилва контактите между представителите на различни нации.

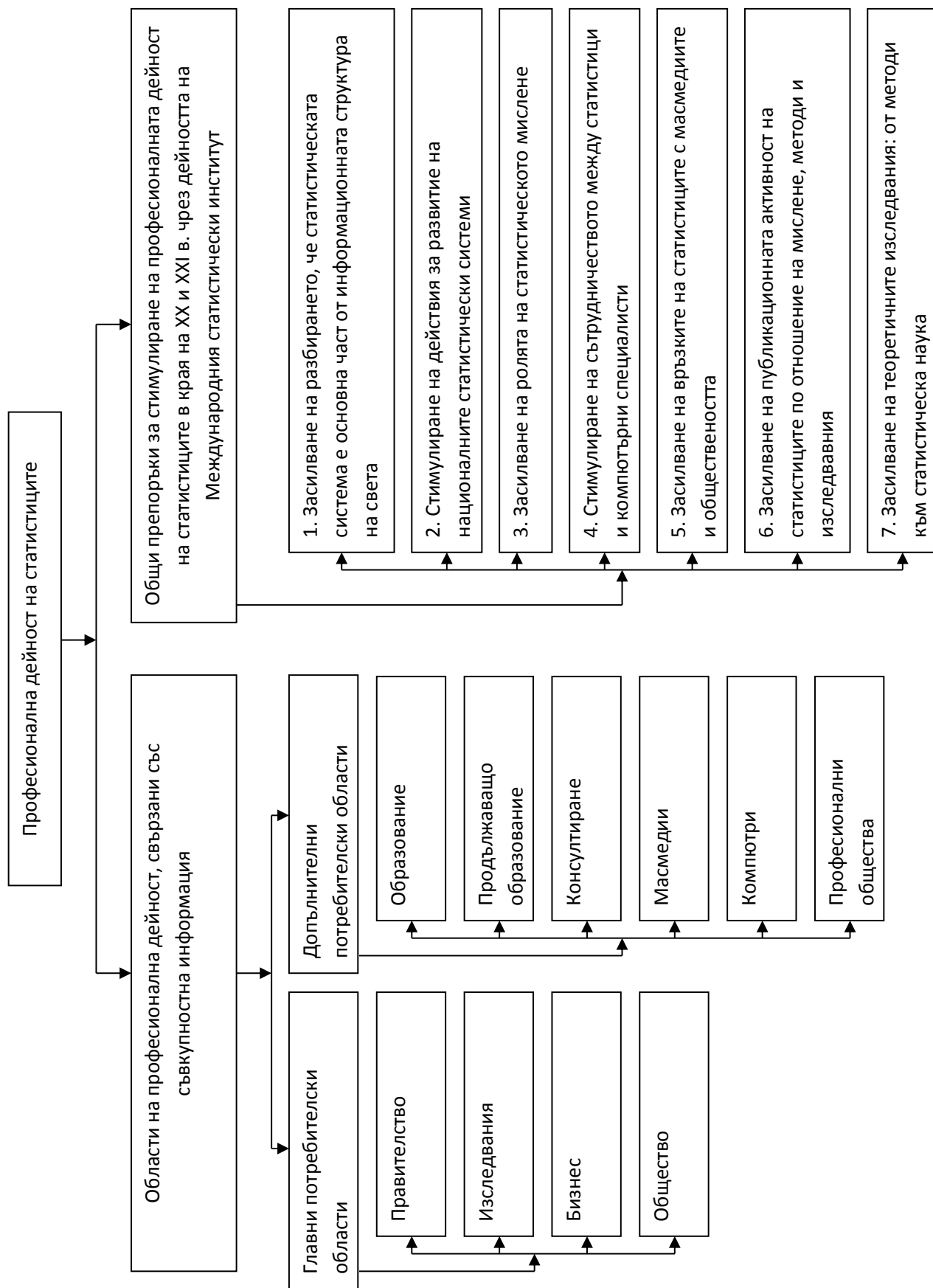
Периодично могат да се изпращат въпросници на бивши официални представители за размерите на статистическия им бюджет, съпоставен с общите правителствени разходи. Резултатите биха били полезни и за сравняването на относителните нива на активност в различните страни дори когато те са експериментални” (Moriguti, S. et al., 1992, pp. 232 - 233).

Препоръка 7. Организацията в рамките на МСИ трябва да предприемат конкретни действия, които следват от изброените общи препоръки, т.е. от методи към статистическа наука. Тук се включва информация, предназначена за обществеността, за политиците, за мениджърите, за ръководителите на изследователски проекти, за общественото мнение, за журналистите, за преподавателите и други с използване на видеокасети и брошури. Различни срещи, в т.ч. семинари, могат да се окажат полезни в съставянето, изпробването и разпространението на тези материали. Ето защо такива срещи заслужават внимание. Комитетите могат да организират дискусии на тема „От статистическите методи към статистическата наука: нови насоки в подготовката на статистиците” (Moriguti, S. et al., 1992, pp. 233 - 234).



Предизвикателствата за промени в съвременния свят в края на XX и през XXI век са предмет на обсъждане на научни конференции, семинари, работни съвещания и в публикации на различни статистически и нестатистически институции, в т.ч. на Международния статистически институт (ISI), на Статистическата служба на европейския съюз (Eurostat), на Конференцията на европейските статистици (CES), на Статистическата комисия и Европейската икономическа комисия на ООН (UNICE), на Организацията за икономическо коопериране и развитие (OECD), на Международния съюз за телекомуникации (ITU) и други.

Фиг. 1. Области на професионална статистическа дейност и препоръки за стимулирането ѝ



МСИ е център, който обединява институционалните структури в отделните страни. Организационната му структура включва: генерална асамблея, съвет, изпълнителен комитет, постоянен офис, асоциации и общества. Асоциациите и обществата са посочени в табл. 1. Прави впечатление, че след 1991 г. има нови международни общества. Те са: Международно общество за бизнес и индустриална статистика и Международно общество за околна среда.

За усъвършенстване на социалния статус, престиж и развиване на мотивацията на статистическата професия заедно с МСИ според утвърдена в практиката традиция участват и други професионални и неправителствени институции: 1) Кралското статистическо общество в Лондон, основано през 1834 г.; 2) Парижкото статистическо общество (1836 г.); 3) Американската асоциация на статистиците, основана през 1839 г.; 4) Международният статистически институт, създаден през 1853 г., и други. В състава на тези професионални съюзи и общества са включени статистици от академичния сектор, бизнеса и статистическите учреждения, които в своето единство са популяризатори на статистическата наука в социалната практика. В България са създадени Съюз на статистиците (през 1990 г.) и Сдружение за статистически изследвания (през 2007 г.) като неправителствени организации.

1. Асоциации и общества в Международния статистически институт

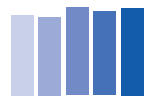
1885: International Statistical Institute	1885: Международен статистически институт
1973: International Association of Survey Statisticians (IASS)	1973: Международна асоциация на изследователите статистици
1975: Bernoulli Society for Mathematical Statistics and Probability (BSMSP)	1975: Общество на Бернули за математическа статистика и вероятности
1977: International Association for Statistical Computing (IASC)	1977: Международна асоциация за компютърни изчисления
1985: International Association for Official Statistics (IAOS)	1985: Международна асоциация на официалната статистика
1991: International Association for Statistical Education (IASE)	1991: Международна асоциация за статистическо образование
2005: International Society for Business and Industrial Statistics (ISBIS)	2005: Международно общество за бизнес и индустриална статистика
2008: The International Environmetrics Society (IES)	2008: Международно общество за околна среда

5. Професионален облик на статистиците

Развитието на съвременното общество отправя предизвикателства за промени в мястото на статистическата наука в статистическата професия и нейния социален статус. Научни дискусии за това се водят на организирания научен семинар от икономическата комисия на ООН за Европа (ЕСЕ) в Лисабон (Португалия) от 25 до 27 септември 1996 г. на тема: „Официалната статистика: минало и бъдеще“. В тях се отбелязва, че „официалната статистика е ориентиране главно към бъдещето, особено във време на социални, икономически и политически промени. Заедно с това е необходим и анализ на миналото, доколкото той позволява да се разбере днешният ден като резултат от историческия процес, да се осъзнаят причините и еволюцията на съвременните явления, да се извлекат уроци от миналото за решаване на съвременните проблеми...“ (Хаглен, Й., 1996, с. 2).

Същевременно научните обсъждания на семинара имат за цел да открият метаморфозите в професията на статистиците като администратори и учени, като практики и теоретици. Основната част от статистиците според Ален Деръозер работят в официалната статистика (в държавните статистически органи). Тя е историческото ядро на професията. Дейността им води своето начало от XVIII век. Същевременно от началото на XX век се разширява професионалната сфера на статистиците. Те работят вече в науката, икономиката, здравеопазването и гражданското общество. В рамките на тези тенденции съществуват известни различия и нюанси в използваните инструменти, традиции и култура на статистиците от различните страни. Това проличава от дейността по създаване на единно европейско пространство и свързаните с нея усилия за сближаване и съгласуване. Създаването на единна институционална структура на Европейския съюз предполага разкриване и съгласуване на противоречията при упражняване на статистическата професия. Това изисква отговор на много въпроси, част от които са: 1) какви са били статистиците в миналото; 2) каква е била тяхната култура и мотивация; 3) как са участвали в различните научни школи; 4) какво е било взаимното им влияние; 5) съществува ли тенденция към засилване на интеграционните функции на статистическата професия в днешно време; 6) какво е бъдещето на статистическата професия (Деръозер, Ал., 1996, с. 1 - 20).

Отговорът на тези въпроси не е и не може да бъде еднозначен. Той е свързан с бавната интеграция на две съвършено различни по своя произход традиции. Те са немската и англосаксонската. Първата традиция се развива като административна на



основата на държавознанието на професорите Херман Конринг и Готфрид Ахенвал (staatenkunde). Тя в по-голяма степен се приближава до таксономията, която има за цел систематизирана класификация на забележителности за държавата и правото, отколкото до метрологията. Втората традиция се основава на политическата аритметика на Дж. Граунт и У. Петти, на естествените науки, в т.ч. на биометрията и на теорията на измерванията и изчисляването на вероятности. Това класическо противопоставяне на немската и англосаксонската традиция през XVIII век в статистиката като наука поражда два основни професионални типа статистици. Едните са държавни служители (civil servants), които изпълняват своите функции в съответствие със законите, нормите и обичаите на държавата, а другите - академични учени (professional), притежаващи специални знания по статистика, прилагани в научни изследвания. Бъдещето на професията на статистиците през XXI век според мен е изправено пред сериозни предизвикателства. Те изискват ускоряване на интеграцията на немската и англосаксонската традиция в професията и промени в развитието на статистиката като наука в условията на съвременното общество, глобализацията и превръщането на знанията в непосредствена производителна сила. В тях статистиката ще бъде същностна част на информационната им инфраструктура. Тя трябва да се развива като система, която да обхване целия свят.

За първи път идеята за провеждане на извадкови изследвания е издигната от Пиер Лаплас, но скоро е отхвърлена и забравена. По-късно тя е възродена от Киер - ръководител на статистическото управление в Норвегия, и подробно описана от проф. А. Боули. Мащабите на използване на този статистически метод остават твърде ограничени. През 1925 г. на заседанията, организирани от Международния статистически институт, участниците не са успели да стигнат до единно мнение по отношение на преимуществата на метода на „случайния подбор“ в сравнение с метода на „щателния подбор“, който е първообраз на квотния метод. През 1934 г. проф. Е. Нейман разработва теорията на стратифицираната извадка. Методите на извадковите изследвания получават широка известност във връзка с проведените проучвания през 1936 г. от Галъп на общественото мнение за американските президентски избори. В България знанията са въведени като практика от проф. О. Андерсон през 1929 г., а като знания в образованието - от проф. В. Цонев през 1958 година.

В същото време в Бюрото за преброяванията в САЩ постъпват на работа младите университетски преподаватели Дедрик и Хансен. Те се занимават с експериментални проучвания за приложенията на извадковия метод при изучаване на

безработицата и на домакинските бюджети. Методът на извадковите изучавания на съвременния етап е основен за производството на статистическа информация в условията на пазарното стопанство. Той е предизвикателство към професионалните знания на статистиците (Деръзер, А., 1996, с. 15 - 19). Още повече, че в настоящия етап няма единна национална политика за събиране на данни от респондентите и използването на административни източници за статистически цели, както и за преминаване от изчерпателни към извадкови годишни изследвания.

В НСИ (България) все още няма специализирано структурно звено за обща методология, планиране, координиране и съвместяване на извадките. Събирането, обработката и разпространението на статистическите данни от извадковите изследвания все още не се провежда онлайн, което е в противоречие с изискванията на съвременното общество.

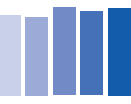
Нови предизвикателства към професионалните статистици възникват след приемането на България за член на Европейската общност. Те са насочени преди всичко към развитието на краткосрочната бизнес статистика, в т.ч. към усъвършенстване на системата от индекси на цените, систематично провеждане на изследвания за използване на ИКТ в предприятията и домакинствата, подготовка на методологии и методики за провеждане на нови изследвания за използване на ИКТ във финансовите предприятия и за инвестициите на предприятията в ИКТ.

6. Опорни точки на статистическата професия и промени в длъжностите

Съществуването и развитието на професията на статистиците при съвременни условия се основава най-малко на пет опорни точки. Те са: 1) Закон за статистиката; 2) Национална класификация на професиите и длъжностите в Република България; 3) съществуването на професия „статистик“ в международната и европейската статистическа практика; 4) съществуването на оторизирана правителствена институция за производство на информация и 5) обществената необходимост от статистическата информация и знания (Радилов, Д., 1998, с. 7 - 8).

Законът за статистиката в чл. 2 определя съдържанието на статистическата дейност, която се осъществява при спазване на определени принципи (професионална независимост, безпристрастност, обективност, надеждност и т.н.) и критерии за качество (адекватност, точност, актуалност, достъпност, яснота, сравнимост и логическа обвързаност).

Националната класификация на професиите и длъжностите, 2011 г. (НКПД-2011) определя професионалната и длъжностната структура в Република България. Тя е



напълно сравнима с Международната стандартна класификация на професиите (ISCO-08).

Професията „статистик“ в международната и европейската статистическа практика е свързана с развитието на широкоспектърната статистика. Тя включва макроикономическа статистика, бизнес статистика, демографска статистика, статистика на социалните дейности, статистика в образованието и здравеопазването. Развитието на съвременния статистически софтуер разширява сферата на приложните статистически изследвания в природните, медицинските, техническите, аграрните и ветеринарномедицинските науки, в хуманитарните и правните науки и в разнообразните социални услуги.

Обществената потребност от професията на статистиците в България е свързана с производството на информация и знания за управлението, както и за приложни изследвания в другите науки. **Статистиците могат да работят като държавни експерти и ръководители и изпълнители на различни нива на длъжности**, посочени в НКПД-2011 (табл. 2), както и на свободна практика като ръководители или членове на разнообразни изследователски екипи.

В НКПД-2011 при дефиниране на обхвата на позициите са отразени изискванията на европейското и националното законодателство, специфични особености, свързани с образователната система и организацията на производството, труда, управлението, заплащането на труда, социалното осигуряване и други. **Длъжността се определя** като съвкупност от функции и задачи, които едно лице изпълнява на работното си място, включително в качеството си на работодател и самонает. **Професията е съвкупност от длъжности**, чиито основни функции и задачи се характеризират с висока степен на сходство. **Специалността е съвкупност от знания и умения за конкретен вид трудова дейност, включена в една професия** и придобивана в резултат на професионално образование и обучение в акредитирани висши училища. Тя е свързана с образователното и квалификационното ниво като съвкупност от знания и умения, необходими за изпълнение на определена длъжност, съобразена с българското законодателство.

НКПД-2011 е разработена в десет класа. Професията на статистиците е в клас 2 „Специалисти“. Те обогатяват съществуващите практико-приложни знания и умения в дадена област. Прилагат научни методи, теории и системно изучават водещи постижения. Компетентното упражняване на повечето от професиите в този клас изисква минимум шесто образователно и квалификационно ниво (ОКС „бакалавър“).

Специалистите изпълняват следните задачи: а) **разработват концепции**, теории, оперативни методи и методики; б) **прилагат съществуващите знания** в областта на физическите, математическите и инженерните науки, информационните и комуникационните технологии, хуманитарните и обществените науки, медицинските и здравните науки; в) **провеждат изследвания** и анализират резултатите от тях; г) обучават лица със специални образователни потребности; д) **подготвят научни доклади** и отчети.

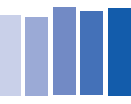
Втори клас (клас 2) включва следните подкласове: **21 „Специалисти по природни и технически науки“**; 22 „Медицински специалисти“; 23 „Преподаватели“; 24 „Стопански и административни специалисти“; 25 „Специалисти по информационни и комуникационни технологии“ и 26 „Юристи и специалисти по обществени науки и култура“.

Специалистите по природни и технически науки провеждат изследвания, усъвършенстват или разработват концепции, теории, оперативни методи и прилагат знания в други области като физика, астрономия, метеорология, химия, геофизика, геология, статистика и други. Те имат следните основни задачи: 1) провеждат самостоятелни изследвания; 2) увеличават или прилагат научни знания, получени чрез изучаване на структури, свойства и процеси на физични явления, химични свойства и всички форми на човешки, животински и растителни видове; 3) консултират, проектират и ръководят строителство на сгради, инфраструктури и други и 4) изготвят научни доклади и отчети.

В този подклас се включват с код 212 „Математици, актюери и статистици“. Те произвеждат научни изследвания и експериментални разработки, усъвършенстват или разработват математически, актюерски или статистически концепции, теории, техники и прилагат тези знания в инженерни, стопански, социални и други науки.

В единичната група с код 2120 се включват посочените в табл. 1 длъжности на статистиците. Те са: 1) биометрик; 2) демограф; 3) статистик; 4) статистик, икономическа статистика; 5) статистик, математическа статистика и 6) статистик, иконометрия.

Необходимо е да се отбележи, че в групата 212 „Математици, статистици и сродни на тях специалности“ за 2011 г. са посочени 5 длъжности, свързани с професията статистик, докато през 2005 г. тези позиции са били 19.



Считам, че включването на професията на статистиците в НКПД-2011 в подклас 21 „Специалисти в областта на физическите, математическите и инженерните специалности“ не е много сполучливо.

Бъдещето на статистическата професия в съвременния свят изисква бързи промени в нейното съдържание в триъгълника: знания - образование - бизнес. Тези промени са свързани с някои предизвикателства към професията. Те са следните: 1) организация на сегменти за производство на статистическа информация и знания, получени от нея; 2) разпространение на статистическа информация и знания; 3) обслужване на потребители и работа за повишаване на тяхната статистическа грамотност; 4) конфиденциалност и прозрачност на статистическите знания и 5) превръщане на статистическите знания в иновации, насочени към бизнеса.

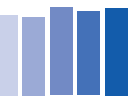
Статистиците могат да работят в екип или индивидуално. Работната им среда може да бъде в офис или извън него, а необходимите им работни инструменти включват компютър, принтер и друг специализиран хардуер и софтуер. **Използването на нови ИКТ пренася професионалната дейност на статистиците във виртуални офиси** (Илиев, П. и др., 2010, с. 8 - 18).

**2. Промени в Националната класификация на професиите (НКП) - 1996 г., и
Националните класификации на професиите и длъжностите (НКПД) -
2005 и 2011 година**

Код НКП-1996	Професии и длъжности	Код НКПД-2005	Код НКПД-2011
2.	Клас 2. Аналитични специалисти		Специалисти
21.	Физици, математици и инженерни специалности		
212.	Математици, статистици и сродни на тях специалности		
	Математици и приравнени на тях		
2121.	Актьор	2121 7001	2120 6001
2121.	Аналитик, изследване на операциите	2121 7002	2120 6002
2121.	Математик	2121 7003	2120 6003
2121.	Математик, математически анализ	2121 7004	
2121.	Математик, приложна математика	2124 7005	
	Статистици		
2122.	Биометрик	2121 7001	2120 6004
2122.	Демограф	2121 7002	2120 6005
2122.	Статистик	2121 7003	2120 6006
2122.	Статистик, биометрик	2121 7004	
2122.	Статистик, биостатистик	2121 7005	
2122.	Статистик, демограф	2121 7006	
2122.	Статистик, изучаване на общественото мнение	2121 7007	
2122.	Статистик, икономическа статистика	2121 7008	2120 6007
2122.	Статистик, математическа статистика	2121 7009	2120 6008
2122.	Статистик, медицинска статистика	2121 7010	
2122.	Статистик, приложна статистика	2121 7011	
2122.	Статистик, промишлена статистика	2121 7012	
2122.	Статистик, селскостопанска статистика	2121 7013	
2122.	Статистик, статистика на образованието	2121 7014	
2122.	Статистик, статистика на транспорта и комуникациите	2121 7015	
2122.	Статистик, трудова статистика	2121 7016	
2122.	Статистик, търговска статистика	2121 7017	
2122.	Статистик, физически науки	2121 7018	
2122.	Статистик, финансова статистика	2121 7019	
	Статистик, иконометрия		2631 6003

7. Промени в образованието и етиката на статистическата професия

Промените в статистическото образование произтичат от включването ни в европейското образователно пространство. То се характеризира със свободно движение на хора, образователни програми, академично признаване на квалификации и продължаващо образование през целия живот. Бъдещето на статистическото образование е изправено пред сериозни предизвикателства за промени. Те произтичат



от промените в съвременния свят, породени от новите информационни технологии, информацията и знанията, от една страна, а от друга, от европейските реформи във висшето образование. Началото на тези промени е поставено през 1999 г. с подписване на Болонската декларация. В нея през 2010 г. се отбелязват шест краткосрочни цели. Те са: 1) възприемане на разбираема и съпоставима система на висше образование с въвеждане на европейско дипломно приложение; 2) възприемане на двустепенно и тристепенно образование (бакалавър, магистър и доктор); 3) въвеждане на система за натрупване и трансфер на кредити с цел засилване на студентската мобилност; 4) засилване на европейските измерения във висшето образование при разработване на учебни програми, междуинституционално сътрудничество, схеми за мобилност и интегриране за обучение и подготовка; 5) засилване на сътрудничеството в областта на оценяване на качеството на висшето образование чрез разработване на съпоставими критерии, методологии и т.н.

В рамките на тази статия не е възможно да се обхванат изчерпателно всички необходими промени в образованието по статистика и най-вече промените в учебните планове за ОКС „бакалавър“ и „магистър“. Като задължителен образователен минимум биха могли да се включат: 1) знания за управление на научни проекти; 2) концептуални знания за създаване на статистически информационни системи и свързания с тях дизайн и експерименти; 3) аналитични знания за количествени инструменти, главно за специализиран статистически и иконометричен софтуер; 4) знания за офис приложения и администриране на компютърни мрежи; 5) знания за нови информационни технологии, например за Big Data, и 6) знания за ръководство и управление на експерименталната дейност на научни колективи и други. Тези знания трябва да са в унисон с перманентното проучване на потребителски профили и маркетинга на статистическата информация.

Статистическото образование трябва да се свърже с електронното обучение. То се характеризира с приемане на разнообразни подходи, методи и инструменти, стандарти за управление на знания, сертификати и инструменти за контрол на качеството на статистическото образование.

Етиката в статистическата професия е система от норми за нравствено поведение, нравствени правила и морал. Етични са ценностите, които определят кое действие на професионалните статистици е добро, правилно или неправилно от морална гледна точка.

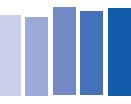
В съвременния свят етичните принципи на статистическата професия са насочени към повишаване на доверието на клиенти, работодатели, политици, журналисти и на всички други потребители от съвременния свят като цяло към статистическите знания. Етичните принципи се разработват от различни професионални неправителствени асоциации и институции. За етичните принципи на статистиката са създадени следните основни документи: 1) Фундаментални принципи на официалната статистика от Статистическия департамент на ООН през 1994 г.; 2) Декларация на Международния статистически институт за статистическа етика през 1985 г.; 3) Ръководство по етика на статистическата практика от Американската статистическа асоциация през 1999 г.; 4) Кодекс на европейската статистическа практика, разработен и допълнен от Евростат през 2011 г., и други.

Етичните принципи, разработени от Американската статистическа асоциация (Ethical Guidelines for Statistical Practice, 1999, pp. 1 - 11), включват осем тематични области. Те са: 1) професионализъм; 2) отговорност към финансиращи лица, клиенти и служители; 3) отговорност при публикации и становища; 4) отговорност за защита на интересите на обектите на статистическите проучвания; 5) отговорност към другите статистици; 6) отговорност при обвинения за неправомерна дейност; 7) отговорност към колегите от изследователския екип и 8) отговорности към работодателите и други лица, които използват статистическата информация и знания.

В Кодекса на европейската статистическа практика (Кодекс на европейската статистическа практика, 2011, с. 1 - 15) се разглеждат 15 принципа в три области. Първата област е институционалната среда. В нея се разглеждат институционални и организационни фактори, оказващи влияние върху ефективността и надеждността на статистическите знания, които се разработват, изготвят и разпространяват; **втората област** са статистическите процеси, които изискват стриктно спазване на европейските и други международни стандарти, насоки и добри практики, и **третата област** са статистическите продукти, които трябва да отговарят на европейските стандарти за качество и да удовлетворяват потребностите на европейските институции, правителства, научни изследователски институти, предприятия и на гражданското общество.

Заключение

Разгледаните академични научни позиции за статистическата наука в съвременния свят позволяват да се формулират следните обобщения: 1) развитието на предмета на статистическата наука като клон на математиката все още не е изживяно; 2)



липсата на дефиниция за предмета на статистическата наука пречи за развитието ѝ в съвременния свят; 3) разграничаването на статистиката от математиката се основава на единството между „статистическата съвкупност” и „множеството” като две страни на единното битие на масовите явления и процеси - материално-веществена и абстрактна; 4) развитието на статистическата наука като социална система от знания трябва да се основава на забравените изследвания на Ламберт Адолф Кетле, Георг фон Майер, Франц Джиджек, Уилям Петти, Оскар Андерсон, Иван Стефанов, М. Ханзен и др.; 5) бъдещето на статистическата наука е като социална система, която включва информация и знания не само през XXI век, но и след него.

Промените в съвременния свят, в който централно място заемат информацията, знанията и комуникациите, са предизвикателства за промени в статистическата професия и свързаните с нея длъжности на трудовия пазар. Те оказват влияние върху опорните точки на статистическата професия и водят до промени в длъжностите, свързани с нея. Необходимо е да се отбележи, че в резултат от промени на трудовия пазар длъжностите: 1) статистик; 2) биометрик; 3) статистик; 4) биостатистик; 5) статистик, демограф; 6) статистик, изучаващ общественото мнение; 7) статистик, медицинска статистика; 8) статистик, приложна статистика; 9) статистик, промишлена статистика; 10) статистик, образователна статистика, и други - общо 14, не фигурират в НКПД-2011. Настъпват промени в професионалния облик на статистиците, свързани с новите информационни технологии. Възникват виртуални офиси, които се отразяват на статистическото образование. В него трябва да намерят място нови знания.

Бъдещето на статистическата професия изисква неотложни промени в учебните планове за получаване на образователно-квалификационни степени „бакалавър” и „магистър”. В тях трябва да бъдат включени съвременни извадкови изследвания със софтуер, както и нови информационни технологии, свързани с Big Data, и т.н. Важно значение има въвеждането на учебна дисциплина за етиката на статистическата професия при нейното упражняване. Това е широка област за контрол на дейността на професионалните статистици, която се основава на етични принципи, разработени и обобщени в Кодекса на европейската статистическа практика. Спазването им би осигурило доверие към статистическите знания и статистическата професия в съвременния свят.

ЦИТИРАНА ЛИТЕРАТУРА:

Александрова, Н. В. (1978). Математические термины. Москва, Высшая школа, с. 79 - 80.

Андерсон, О. (1928). За репрезентативния метод и неговото приложение при разработката на материалите от преброяване на земеделските стопанства на 31.XII.1926 г. - Тримесечно списание на Главната дирекция на статистиката. София, кн. 2 и 3, с. 109 - 152.

Българска енциклопедия (2002). Второ преработено и разширено издание, БАН, Книгоиздателска къща Труд, С., с. 1331.

Дерозьер, А. (1996). Администратор и ученый: метаморфозы, профессии статистика. Семинар по официальной статистики - прошлое и будущее, Лиссабон, Португалия, 25 - 27 септември.

Данаилов, Г. Т. (1932). Теория на статистиката. София, Държавна печатница, с. 1 - 638.

Димитров, Б. и Н. Янев (1998). Вероятности и статистика. Унив. изд. „Св. Климент Охридски“, София, с. 167 - 173.

Димитров, Д. (2004). Теория на вероятностите и математическа статистика. Изд. „Наука и икономика“, Икономически университет - Варна, с. 16.

Илиев, П. и др. (2010). Виртуални системи. Библиотека „Проф. Цани Калянджиев“, изд. „Наука и икономика“, Икономически университет - Варна, с. 8 - 18.

Кендалл, М., А. Стюарт (1966). Теория распределений. Перев. с англ. Изд. „Наука“. Глав. ред. Физ.-мат. лит. Москва, с. 1 - 587.

Кетле, А. (1836). Социальная статистика или опыт исследования о развитии человеческих способностей. Перев. с англ. 1911, Киев, т. 1, с. 16 - 17.

Кодекс на Европейската статистическа практика за националните и общностните статистически органи, приет от Комитета на Европейската статистическа система на 28.09.2011 г. Европейска статистическа система, Eurostat, с. 1 - 15.

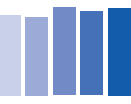
Майер, фон Г. (1899). Статистика и обществоведение, том първий, Теоретическая статистика. Пер. с нем. приват-доц. В. Я. Железнова. С. Петербургъ. Типография товар. „Народная польза“, с. 5 - 221.

Марчев, Ф. (2003). Философия: идеи, школи, направления - 2-ро доп. и прераб. изд. „Варна принт“, Варна.

Национална класификация на професиите и длъжностите (2005), в сила от 1.01.2006 г., ИК „Труд и право“, София.

Национална класификация на професиите и длъжностите (2011). <http://www.mlsp.government.bg/class1/store/listclass.asp>.

Нови схващания върху теоретичната статистика (1940). Академична реч на проф. Иван Стефанов в гр. Свищов на 15 май 1940 г. на тържествено годишно събрание на Висшето търговско училище „Димитър А. Ценов“ - Свищов. Година IV 1939 - 1940. Печатница „Одеон“, София, с. 37 - 53.



Петканчин, Б. (1968). Основи на математиката. Изд. „Наука и изкуство”, София, с. 5 - 6, 13.

Плошко, Б. Г., И. И. Елисеева (1990). История статистики. „Финансы и статистика”, Москва, с. 1 - 295.

Радилов, Д. (1998). Професията на статистиците в управлението: промени и развитие. Статистика, кн. 6, с. 3 - 15.

Радилов, Д. (2012). Иновативност на статистическите знания във виртуалното пространство. Статистика, информатика, технологии и комуникации. Издателски комплекс на УНСС, София, с. 25 - 30.

Радилов, Д. (2012). Нова парадигма на статистическата наука през ХХІ век. Съвременни методи и технологии в научните изследвания. Сборник с доклади от международна научна конференция. Изд. „Наука и икономика”, Икономически университет - Варна, с. 41, 47.

Радилов, Д. С. Статистическая информация и знания в глобальном мире. Статистические методы в гуманитарных и экономических науках. Материалы международной научно-практической конференции, Санкт-Петербург, 28 - 29 января 2016 г., Санкт-Петербург: Нестор-История, 2016, с. 248 - 251.

Хаглен, Й. (1996). История статистики - проблемы и перспективы международного сотрудничества, Лиссабон, Португалия, 25 - 27 сентября, с. 1 - 21.

Цонев, В. и С. Бадева-Цонева (1983). Етимология, смислово развитие и понятийна изразителност на думата съвкупност. Статистика, кн. 1, с. 13 - 23.

Чолаков, Н. (2007). Трудова и социална статистика. Унив. издателство „Стопанство”, София, с. 437 - 440.

Bowley, A. (1915). Elements of Statistics. London: P. S. King, 1st 1901 - 6th edn, 1937.

Ethical Guidelines for Statistical Practice (1999).

Eurostat, Fields of Education and Training (1999), pp. 3, 18.

Fisher, Nicolas (2001). Critical Issues for Statistics in the Next Two Decades, ISR, vol. 69, 1, pp. 3 - 4.

Kenessey, Z. (1994). Scanning The Future of Statistics, An International Perspective, pp. 1 - 14.

Kenessey, Z. (1994). International Statistical Organization - The Future of Statistics. An International Perspective (1994). Edited by Z. Kenessey, Voorbug, pp. 109 - 146.

Kleene, S. (1952). Introduction to metamathematics, New York. Русский перевод: Клини С., Введение в метаматематику, М., ИЛ, 1957.

Lindley, Denis V. (2000). The philosophy of statistics, Minehead, UK - The Statistician. Royal Statistical Society, 49, part 3, pp. 293 - 318.

Mahalanobis, P. C. (1957). The Foundation of Statistics, Sankhya, 18, pp. 183 - 194.

Moriguti, S., Peter J. Diggle, John C. Gover, Katherine K. Wallman, Shou Ren Wang (1992). The Role of Statisticians, ISR, 60, 3, pp. 227 - 246.

Olenski, J. (2003). The Citizens' Right to Information and the Duties of a Democratic State in Modern IT Environment // ISR, 1, pp. 33 - 48.



Prepared by the Committee on Professional Ethics (</committees/commdetails.cfm?txtComm=CCNPRO03>).

The Oxford Dictionary of Statistical Terms (2003). The International Statistical Institute, edited by Yadolah Dodge, OXFORD University press, p. 388.

Vere-Jones, D. (1995). The coming of age of Statistical education - International Statistical Review, 44, 2.

АКАДЕМИЧНИ НАУЧНИ ПОЗИЦИИ ЗА СТАТИСТИЧЕСКАТА НАУКА И ЗА ПРОМЕНИТЕ В ПРОФЕСИЯТА НА СТАТИСТИЦИТЕ В СЪВРЕМЕННИЯ СВЯТ

*Димитър Радилев**

РЕЗЮМЕ В съвременния свят централно място заемат информацията, знанията и комуникациите. Схващането, че статистиката е клон на математиката, все още не е изживяно. Липсата на дефиниция за предмета на статистическата наука пречи за развитието ѝ в съвременния свят. Разграничаването на статистическата наука от математиката се основава на единството между „статистическата съвкупност“ и на „множеството“ като две страни на единното битие на масовите явления - „материално-веществена“ и „абстрактна“. М. Кендал и А. Стюарт въвеждат понятието „природен феномен“, с което дават много широко определение, включващо всички възможни обекти на външния свят, които са свързани или не са свързани с човека. Развитието на статистическата наука като социална система от знания се основава на забравените изследвания на Ламберт Адолф Кетле (1836), Георг фон Майер (1895 - 1909), Франц Джиджек (1940), Уилям Петти (1691), Оскар Андерсон (1928), Ханзен (1937) и т.н. Мястото на статистическата професия се разглежда за първи път на Общата асамблея на Международния статистически институт през 1987 г. в Токио (Япония). Сформирана е комисия с ръководител проф. Сигети Моригути с широки пълномощия, която трябва да анализира мястото на статистическата професия в съвременния свят. Тя прави 7 препоръки. По-късно, от 25 до 27 септември 1996 г. в Лисабон, Португалия, се провежда семинар за официалната статистика и за нейното минало и бъдеще. Настъпилите промени в професионалния облик на статистиците след навлизането на новите информационни технологии (компютър, интернет, мобилен телефон, сателитни чинии, облачни технологии и др.) водят до съществени промени в живота, поведението и ценностната система на статистиците. Възникват виртуални офиси, които изискват промени в статистическото образование при спазване на етични принципи.

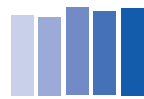
* Проф. д-р ик. н. в катедра „Статистика и приложна математика“, Икономически университет - Варна;
e-mail: d.radilov@gmail.com.

АКАДЕМИЧНЫЕ НАУЧНЫЕ ПОЗИЦИИ В СТАТИСТИЧЕСКОЙ НАУКЕ И О НАСТУПИВШИХ ПЕРЕМЕНАХ В ПРОФЕССИИ „СТАТИСТИК“ В СОВРЕМЕННОМ МИРЕ

Димитр Радилов*

РЕЗЮМЕ Центральное место в современном мире занимают информация, знания и коммуникации. Точка зрения о том, что статистика является разделом математики, все еще остается актуальной. Отсутствие определения предмета статистической науки препятствует ее развитию в современном мире. Разграничение статистической науки от математики основывается на единстве „статистической совокупности» и „множества», как двух сторон существования массовых явлений - „материально-вещественной» и „абстрактной». М. Кендалл и А. Стюарт вводят понятие „природный феномен», которое включает в довольно обширное определение все возможные объекты внешнего мира, которые связаны или не связаны с человеком. Развитие статистической науки, как социальной системы знаний, основана на забытых исследованиях Ламберта Кетле (1836), Георга фон Майера (1895-1909), Франца Джиджека (1940), Уильяма Петти (1691), Оскара Андерсона (1928), Ханзена (1937) и т. д. Впервые вопрос о месте статистической профессии обсуждался на Генеральной Ассамблее Международного института статистики, которая состоялась в Токио (Япония) в 1987 году. Была создана комиссия, под руководством профессора С. Моригути, которая получила большие полномочия для анализа места статистической профессии в современном мире. Комиссия изготовила 7 рекомендаций. Позже, с 25 по 27 сентября 1996 года в Лиссабоне (Португалия), состоялся семинар по официальной статистике, о ее прошлом и будущем. С появлением новых информационных технологий (компьютера, интернета, мобильного телефона, спутниковой тарелки, облачных технологий и т. д.) наступили изменения в профессиональном плане статистиков, что довело до существенных перемен в их жизни, поведении и в системе ценностей. С появлением виртуальных офисов, возникает необходимость изменений в статистическом образовании в соответствии с соблюдениями этических принципов.

* Д-р экон. наук, проф. кафедры „Статистика и прикладная математика“, Экономический университет - Варна; e-mail: d.radilov@gmail.com.



**ACADEMIC SCIENTIFIC POSITIONS ON STATISTICAL SCIENCE AND ABOUT
CHANGES IN THE PROFESSION OF STATISTICIANS IN
THE CONTEMPORARY WORLD**

*Dimitar Radilov**

SUMMARY In contemporary world information, knowledge and communications have their central position. The assumption that statistics is a branch of mathematics has not yet been outlived. The lack of a definition about the subject of statistical science hinders its development in the modern world. The distinction between statistical science and mathematics is based on the unity between the ‘statistical aggregate’ and the ‘multiplicity’ as two sides of the unity of mass phenomena - ‘material and substantial’ and ‘abstract’.

M. Kendal and A. Stuart introduce the term of a ‘natural phenomenon’, with which they provide a very broad definition, including all the possible objects of the outer world that are related or unrelated to man. The development of statistical science as a social system of knowledge is based on the forgotten studies of Lambert Adolf Kettle (1836), Georg von Meyer (1895-1909), Franz Djidjeck (1940), William Petty (1691), Oscar Anderson (1928), Hanzen (1937) and so on. The place of the statistical profession is first considered in the 1987 General Assembly of the International Institute of Statistics in Tokyo (Japan).

A commission is established, headed by Prof. Sigeiti Moriguti, with a broad mandate to analyse the place of the statistical profession in the modern world. It makes seven recommendations. Later, from 25 to 27 September 1996 in Lisbon, Portugal, a seminar on official statistics and on its past and future took place. The changes occurring in the professional image of the statisticians after the introduction of the new information technologies (computer, internet, mobile phone, satellite dishes, cloud technologies, etc.) lead to significant changes in the life, behaviour and value system of the statisticians. Virtual offices are emerging, requiring changes in statistical education in line with ethical principles.

* Prof. Dr. in the Department of Statistics and Applied Mathematics, University of Economics - Varna; e-mail: d.radilov@gmail.com.

КЛЪСТЕРИЗАЦИЯ НА ЕВРОПЕЙСКИТЕ СТРАНИ СПОРЕД СТЕПЕНТА НА ТУРИСТИЧЕСКА КОНКУРЕНТОСПОСОБНОСТ

*Тихомир Върбанов**



Въведение

Туристическият сектор е един от водещите и най-динамично развиващи се сектори в световната икономика, който оказва въздействие за развитието на редица държави не само в икономически, но и в социален и екологичен аспект. За много страни той е основен източник на инвестиции и приходи, като икономическите ползи от туристическа дейност се разпространяват и сред редица други сектори, с които той е тясно свързан - търговия на дребно, селско стопанство, транспорт, строителство и т.н. Поради това туризмът има не само пряко, но и косвено значение за повишаването на брутният вътрешен продукт както на национално, така и на световно равнище.

Независимо от глобалната геополитическа несигурност и икономическа нестабилност през последното десетилетие туристическият сектор запазва своето устойчиво развитие. Според доклада на Световния съвет за пътуване и туризъм (World Travel and Tourism Council - WTTC)¹ през 2016 г. за шеста поредна година се отчита нарастване на сектора, в т.ч. и по отношение на броя на международните туристически пътувания. В световен план през 2016 г. прекият принос на сектора към глобалния брутен вътрешен продукт достига 2 306 млрд. долара, или 3.1% в относителен размер, а цялостният възлиза на 7 613.3 млрд. долара (10.2%). Освен това туристическият сектор генерира голям брой работни места, като общият принос към пазара на труда възлиза на 292 млн. работни позиции (т.е. всяко десето заето лице извършва свързана с туризма икономическа дейност), а пряко с туристическа дейност са ангажирани 109 млн. души.

Според изследванията на Световния съвет за пътуване и туризъм в световен мащаб се очаква тенденцията към нарастване на сектора да се запази и в дългосрочен

* Асистент д-р в катедра „Математика и статистика“, СА „Д. А. Ценов“ - Свищов; e-mail: t.varbanov@uni-svishtov.bg.

¹ Travel & Tourism Economic Impact 2017 World. <https://www.wttc.org/>.

план, в т.ч. и по отношение на разкриването на нови работни места, чийто дял от общо създадените във всички икономически сектори ще достигне около една четвърт, а броят на пряко заетите лица в сектора ще надвиши 138 млн. души. Анализите показват², че за създаването на едно работно място в туристическия сектор е необходимо привличането на 30 нови туристи, което извежда на преден план през следващото десетилетие въпросите, свързани с устойчивото развитие и **конкурентоспособността** на туристическите дестинации.

Европейските държави формират регион с традиционно силно развити туристически дестинации поради богатото културно наследство, добре изградената туристическа инфраструктура и високата степен на интеграция. Туризмът има съществена роля за постигане на редица стратегически цели на ЕС, сред които устойчиво икономическо развитие, развитие на човешките ресурси, икономическо и социално сближаване и други. Относителният дял на сектора в общоевропейския брутен вътрешен продукт и заетост през 2016 г. е съответно 3.7 и 5%, което в абсолютна стойност се равнява на 611.3 млрд. долара и 11.4 млн. души. Ако се отчете и индиректният принос, то тогава генерираният БВП и заетост значително нарастват - 1 682.9 млрд. долара (10.2% от общия БВП) и 26.6 млн. души (11.6% от всички заети лица).

Европа е регионът с най-висок пазарен дял в света и през 2016 г. е привлякъл близо 52% от международните туристи. Въпреки това през последните години се отчита известно изоставане в темповете, с които се развива туристическият сектор в Европа в сравнение с регионите с най-бързо развиващ се туризъм - Югоизточна и Южна Азия. Отчитайки това, прогнозата на Световната организация по туризъм (World Tourism Organization - WTO)³ предвижда забавяне в прираста на туристическия сектор в Европа до 2030 г., в резултат на което пазарният дял ще се свие до 41.1%. За да запази ролята си на водещ туристически регион, е необходимо да се изследват причините за загубата на пазарни позиции и да се изготвят стратегии за повишаване на конкурентоспособността и устойчиво развитие на туристическия сектор.

Европейските страни се характеризират с различни равнища на туристическа конкурентоспособност в зависимост от наличието или липсата на определени конкурентни предимства, както и от степента, до която те са развити. **Целта** на настоящата статия е да се извърши клъстеризация на включените в анализа 38 европейски страни в еднородни групи според факторите, определящи конкурентоспособността на туристическия сектор, и да се изследват силните и слабите страни на дестинациите. Информационното осигуряване на анализа се основава на докладите на Световния икономически форум (World Economic Forum - WEF) за конкурентоспособността на пътуванията и туризма (The Travel & Tourism Competitiveness Report: 2015, 2017).

² The Travel & Tourism Competitiveness Report 2015: Growth through Shocks. p. 3. <http://www3.weforum.org>.

³ Tourism towards 2030. Global overview. <http://www2.unwto.org/>.

I. Теоретична рамка на изследването

През последните десетилетия се наблюдава тенденция към подобряване на качеството на предлаганите туристически услуги и значителни инвестиции в туристическа инфраструктура, което е предпоставка все по-голям дял от разполагаемите доходи на домакинствата да се изразходват с цел пътувания и туризъм. Това води до все по-силно изразена конкуренция на световния туристически пазар, тъй като отделните дестинации непрекъснато се стремят към по-големи пазарни дялове. Според редица автори успешното реализиране на дестинацията до голяма степен се определя от достигнатото равнище на конкурентоспособност.

В научната литература са изведени голям брой дефиниции за конкурентоспособност на туристическата дестинация предвид многоаспектния и комплексен характер на категорията. Те се основават на отделни елементи на конкурентоспособността и най-често акцентират върху удовлетвореността на потребителите, икономическия и социалния ефект и други.

Така например **Enright** и **Newton**⁴ определят конкурентоспособността като възможност да бъдат привлечени потенциални туристи и да бъдат задоволявани техните потребности, а **d’Hauteserre**⁵ - като способност на туристическата дестинация да запазва или подобрява своята пазарна позиция спрямо преките си конкуренти.

Hassan⁶, **Chon** и **Meyer**⁷ и **Hong**⁸ разширяват тези концепции, като включват създаване, интегриране и предоставяне на продукти и услуги с добавена стойност, които отговарят на изискванията на световния туристически пазар, повишаване на доходите на местното население и като цяло трансформиране на сравнителните предимства в конкурентни. В резултат се постига устойчиво използване на природните ресурси на дестинацията и се поддържат добри пазарни позиции спрямо преките конкуренти.

Според **Dwyer**⁹ основен фокус трябва да се постави върху пазарните механизми и компонентите на туристическия сектор, които са в основата на облика на дестинацията като цяло. Той счита, че конкурентоспособността на дестинацията трябва да се разглежда като цялостна концепция, която включва обвързаните с валутните курсове различия в цените, равнищата на производителност на различните компоненти на туристическата индустрия и характеристиките на средата, които допринасят за привлекателността на дестинацията сред туристите.

⁴ **Enright, M.**, Newtown, J. Determinants of tourism destination competitiveness in Asia Pacific: Comprehensiveness and universality. // *Journal of Travel Research*, 2005, No. 43, p. 340.

⁵ **D’Hauteserre A. M.** Lessons in managed destination competitiveness: The case of Foxwoods Casino Resort. // *Tourism Management*, 2000, Vol. 21, No. 1, pp. 23 - 32.

⁶ **Hassan, S.** Determinants of market competitiveness in an environmentally sustainable tourism industry. // *Journal of Travel Research*, 2000, No. 38, pp. 240.

⁷ **Chon K. S.**, Mayer K. J. Destination competitiveness models in tourism and their application to Las Vegas. // *Journal of Tourism Systems and Quality Management*, 1995, Vol. 1, No. 2/4, pp. 227 - 246.

⁸ **Hong, S.** Competitiveness in the Tourism Sector, A Comprehensive Approach from Economic and Management Points. Springer Science and Business Media, 2008, p. 6.

⁹ **Dwyer, L. P.**, Forsyth, P., Rao, P. The price competitiveness of travel and tourism: a comparison of 19 destinations. // *Tourism Management*, 2000, No. 21, pp. 9 - 22.

Ritchie и **Crouch**¹⁰ разглеждат конкурентоспособността от гледна точка на дългосрочния икономически просперитет. Определят я като способност на дестинацията да увеличава разходите за туризъм и да привлича все по-голям брой туристи, да задоволява техните изисквания и да им предоставя запомнящи се преживявания. Същевременно с това устойчиво повишава благосъстоянието на своите резиденти и запазва природните и културните ресурси на дестинацията за бъдещите поколения.

Една от **най-широко възприетите дефиниции** е формулирана от Организацията за икономическо сътрудничество и развитие (ОИСР)¹¹. Според нея конкурентоспособността се разглежда като способност на дестинацията да повишава своята привлекателност сред местното население и посетителите, да предоставя качествени, иновативни и икономически рентабилни туристически услуги и да повишава своя дял на световния туристически пазар в условия на ефективно и устойчиво използване на наличните ресурси. В настоящото изследване приемаме като базова за анализа именно тази дефиниция.

В обобщение на изложените концепции можем да посочим, че конкурентоспособността е решаващ фактор за успеха на туристическата дестинация. Нейното повишаване е свързано със запазване на постигнатите и придобиването на нови стратегически конкурентни предимства и се превръща в основна задача в управлението на туристическата дестинация. Това е важно условие за състоянието на туристическия сектор и на икономиката като цяло, поради което е необходимо да се развият конкурентните предимства на дестинацията. За целта трябва да се установи нейната действителна пазарна позиция, което, от една страна, е свързано с точното и навременно установяване на преките конкуренти, а от друга - с определянето на предимствата и недостатъците спрямо тях. Идентифицирането на силните и слабите страни на дестинацията е необходимо условие, без което не би било възможно развитието на туристическия сектор и повишаването на приходите от туризъм.

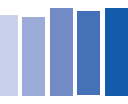
II. Модели на факторната обусловеност на туристическата конкурентоспособност

Ефективното управление на дестинацията е невъзможно без познаване на основните фактори (елементи) на конкурентоспособността. За тяхното извеждане са разработени голям брой теоретични модели, които се основават на широк набор от количествени и качествени индикатори. Най-общо те могат да се систематизират в **две основни групи**. Към първата се отнасят т.нар. **обективни индикатори** за конкурентоспособност, а към втората - **субективните индикатори**, които се основават на възприятията на туристите. Както отбелязва **Crouch**¹², отделните индикатори са с различна степен на важност, поради което вниманието на изследователите трябва да се

¹⁰ **Ritchie J. R. B.**, Crouch G. I. The competitive destination: A sustainable tourism perspective. CABI, Wallingford, 2003, pp. 24 - 26.

¹¹ **Dupeyras, A.**, MacCallum, N. Indicators for Measuring Competitiveness in Tourism: A Guidance Document, OECD Tourism Papers, 2013/02, OECD Publishing, p. 7.

¹² **Crouch, G. I.** Destination Competitiveness: An Analysis of Determinant Attributes. // *Journal of Travel Research*, 2011, Vol. 50 (1), pp. 27 - 45.



насочи единствено към тези фактори, които могат да се използват като основа за вземане на ефективни управленски решения и имат най-голямо значение за повишаване на конкурентоспособността.

Poon¹³ е един от първите автори, които изследват факторната обусловеност на туристическата конкурентоспособност. Извежда следните **четири ключови принципа**, които са в основата на конкурентоспособността на дестинацията:

- акцентирание върху околната среда;
- превръщане на туристическия сектор във водещ за икономиката;
- изграждане на ефективни дистрибуционни канали;
- развитие на частния сектор.

При изследване на конкурентоспособността на единадесет карибски туристически дестинации **De Keyser** и **Vanhove**¹⁴ разработват модел, въз основа на който извеждат пет основни групи фактори. Съгласно резултатите от проведеното изследване от решаващо значение са:

- политиките и стратегиите в областта на туризма;
- макроикономическата среда;
- предлагането;
- транспортната инфраструктура;
- търсенето.

Според **Mihalic**¹⁵ с определящо значение за конкурентоспособността са качеството на околната среда и състоянието на природните ресурси, тъй като те са сред основните мотиви за посещение на дадена туристическа дестинация.

Предложеният модел на **Hassan**¹⁶ се основава на концепцията за сравнителните конкурентни предимства, които авторът разглежда като средство за успешното позициониране на международния туристически пазар. В него околната среда е представена като основна детерминанта, следвана от други три конкурентоопределящи фактора:

- сравнителните предимства;
- отрасловата структура на икономиката;
- търсенето.

В модела на **Yoon**¹⁷ са разгледани вътрешните за дестинацията фактори, които са представени като основа за постигане на конкурентни предимства. Изследвано е както прякото, така и косвеното влияние върху конкурентоспособността на следните три групи фактори:

- привързаност към мястото;
- отношение към околната среда;

¹³ **Poon, A.** Tourism, technology and competitive strategies. CABI, 1993.

¹⁴ **De Keyser, R., Vanhove, N.** The competitive situation of tourism in the Caribbean area methodological approach. // *Revue de Tourisme*, 1994, Vol. 49 (3), pp. 19 - 22.

¹⁵ **Mihalic, T.** Environmental management of a tourist destination A factor of tourism competitiveness. // *Tourism Management*, 2000, No. 21, pp. 65 - 78.

¹⁶ Op. cit., **Hassan, S.** Determinants..., pp. 239 - 245.

¹⁷ **Yoon, Y.** Development of a structural model for tourism destination competitiveness from stakeholders' perspectives (Doctoral dissertation, Virginia Tech), 2000.

- въздействия на туристическото развитие.

Според **Heath**¹⁸ за повишаване на конкурентоспособността на туристическата дестинация значима роля имат множество организации от частния и публичния сектор, като всяка от тях изпълнява специфична функция за постигането на крайната цел. Поради това в своя модел авторът включва следните ключови фактори:

- споделена визия за развитието на дестинацията;
- ясно определени управленски принципи и ценности;
- развитие на човешките ресурси.

Според **Navickas** и **Malakauskaite**¹⁹ конкурентоспособността е от особена важност най-вече за тези дестинации, които искат да увеличат своя дял на динамично развиващия се туристически пазар. Считат, че икономическото развитие и цялостният имидж на дестинацията са от първостепенно значение за тяхната конкурентоспособност, а чрез инвестиции в развитието на туристическия сектор ще се подобри техният международен образ. Според тях страните, които са по-силно развити икономически, разполагат с по-добра инфраструктура, а оттам и с по-конкурентоспособен туристически сектор. За мониторинг на постигнатото равнище на конкурентоспособност на дестинациите авторите предлагат да се използват следните осем групи показатели:

- за оценка на влиянието на туризма за икономическото развитие;
- за ценова конкурентоспособност;
- за развитие на инфраструктурата;
- за оценка на екологията;
- за технологичното развитие;
- за човешките ресурси;
- за степента на отвореност на пазара;
- за социалното развитие.

ОИСР²⁰ предлага да се наблюдават **единадесет ключови индикатора** за оценка на конкурентоспособността, които са обобщени в четири основни категории:

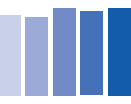
- за оценка на производителността и въздействието на туризма;
- за мониторинг на качеството и конкурентоспособността на предоставяните туристически услуги;
- за оценка на привлекателността на дестинацията;
- за оценка на политическата и икономическата среда.

Конкретните индикатори са: пряк принос на туризма за обема на брутния вътрешен продукт, среден приход от един турист, общ брой нощувки, износ на туристически услуги, производителност на труда в сектора, паритет на покупателната способност и цени, визов режим, природни ресурси и биоразнообразие,

¹⁸ **Heath, E.** Towards a model to enhance destination competitiveness: A Southern African perspective. // *Journal of Hospitality and Tourism Management*, 2002, Vol. 10 (2), pp. 124 - 141.

¹⁹ **Navickas, V., Malakauskaite, A.** The Possibilities for the Identification and Evaluation of Tourism Sector Competitiveness Factors. // *Engineering Economics*, 2009, Vol. 61(1), pp. 37 - 44.

²⁰ Op. cit., **Dupeyras, A.** Indicators ..., pp. 7 - 8.



удовлетвореност на посетителите, културни ресурси и национални стратегии за развитие на туризма.

С най-широко приложение в практиката са теоретичните концепции на **Ritchie** и **Crouch** и **Dwyer** и **Kim**. Методологичната основа на тези модели е концепцията на **M. Porter** за конкурентното предимство на нациите, или т.нар. „диамант на Портър“, която е адаптирана към специфичните особености на туристическия сектор. Както отбелязват **De Holan** и **Phillips**²¹, при изследванията на туристическата конкурентоспособност е препоръчително да се използва именно тази концептуална рамка особено в случаите, когато се изследват развиващи се дестинации.

В разработения от **Ritchie** и **Crouch**²² концептуален модел са представени факторите, които оказват влияние върху конкурентоспособността и устойчивото развитие на туристическата дестинация, както и взаимодействието между тях. Според авторите на модела между двете категории е налице равенство. Те акцентират върху факта, че дестинацията е конкурентоспособна единствено в случаите, когато нейното развитие е устойчиво както в икономически, така и в екологичен, културен и социален аспект. В своя модел включват пет основни групи конкурентоопределящи фактори:

- основни ресурси и атрактори;
- поддържащи фактори и ресурси;
- политика, планиране и развитие на дестинацията;
- управление на дестинацията;
- модифициращи и допълващи фактори.

Авторите считат, че е необходимо да се прави разграничение между конкурентно и сравнително предимство. В обхвата на **сравнителните предимства** се включват всички основни ресурси на дестинацията (наследени и създадени), към които се отнасят местоположението на дестинацията, човешките, физическите и капиталовите ресурси, природните, историческите и културните забележителности, туристическата инфра- и суперструктура и други. **Конкурентните предимства** се свързват с ефективното използване на основните ресурси в дългосрочен план и са основен фактор за привличане на потенциални туристи.

В разработения от **Dwyer** и **Kim**²³ „интегриран модел за анализ на конкурентоспособността на дестинацията“ са посочени шест основни групи фактори:

- наследени ресурси;
- създадени ресурси;
- поддържащи фактори и ресурси;
- управление на дестинацията;
- ситуационни условия (условия на средата);

²¹ **De Holan, P. M., Phillips.** Sun, sand, and hard currency: Tourism in Cuba. // *Annals of Tourism Research*, 1997, Vol. 24 (4), p. 781.

²² **Ritchie J. R. B., Crouch G. I.** Tourism, competitiveness, and societal prosperity. // *Journal of Business Research*, 1999, Vol. 44 (3), pp. 137 - 152; **Ritchie J. R. B., Crouch G. I.** The competitive destination: A sustainability perspective. // *Tourism Management*, 2000, Vol. 21 (1), pp. 1 - 7; **Ritchie J. R. B., Crouch G. I.** The competitive destination: A sustainable tourism perspective, CAB International, Wallingford, 2003.

²³ **Dwyer, L., Kim, C.** Destination competitiveness: determinants and indicators. // *Current issues in tourism*, 2003, Vol. 6 (5), pp. 369 - 414.

- търсене.

Наследените (културно наследство и природни ресурси), създадените (туристическа инфраструктура и атракциони) и поддържащите фактори и ресурси (общата инфраструктура, достъпността, качеството на предоставяните услуги, гостоприемството на местното население и др.) обхващат тези характеристики на дестинацията, които я правят привлекателна за посещения от туристите. Дейностите по управление, планиране и развитие на дестинацията, както и човешките ресурси и околната среда, са включени в обхвата на фактора „управление на дестинацията“. Ситуационните условия включват географското местоположение на дестинацията, микро- и макросредата, сигурността, безопасността и ценовата конкурентоспособност, а условията на търсенето - елементите възприятие, предпочитание и интерес.

Представените теоретични концепции намират широко приложение в практиката - както в емпиричните изследвания върху отделни дестинации и региони, така и за оценка на конкурентоспособността на туристическия сектор в световен план. Световният съвет за пътуване и туризъм и Световният икономически форум периодично публикуват доклади, в които представят резултати относно конкурентоспособността, предимствата и недостатъците на сектора на национално и регионално ниво. Независимо от съществуващите критики към отделни аспекти на използваните методологични рамки²⁴ резултатите от тези изследвания са от изключителна важност за развитието на туризма.

III. Клъстеризация на европейските страни според туристическата конкурентоспособност

Една от най-издържаните методики за оценка на туристическата конкурентоспособност се прилага от Световния икономически форум при изчисляването на индекса на конкурентоспособност на пътуванията и туризма (Travel & Tourism Competitiveness Index - TTCI)²⁵. Индексът се изчислява за повече от 130 страни въз основа на широк набор от количествени и качествени променливи. Чрез тях се оценяват елементите, осигуряващи развитието на туристическия сектор в отделните държави. Информационната осигуреност включва публично достъпни данни, специално организирано изследване сред експерти и организации от туристическия сектор (Executive Opinion Survey), предоставени данни от партниращи изследователски институти и бизнес организации²⁶.

Индексът на конкурентоспособността на пътуванията и туризма е съставен от четири субиндекса:

- „Външна среда“;
- „Туристическа политика“;
- „Инфраструктура“;

²⁴ Mazanec, J. A., Ring, A. Tourism destination competitiveness: second thoughts on the World Economic Forum reports. // *Tourism Economics*, 2011, Vol. 17 (4), p. 729; Crouch, G. I. Measuring tourism competitiveness: research, theory and the WEF index. Proceedings of the ANZMAC, 2007, pp. 73 - 79.

²⁵ Op. cit., Dupeyras, A. Indicators..., p. 13.

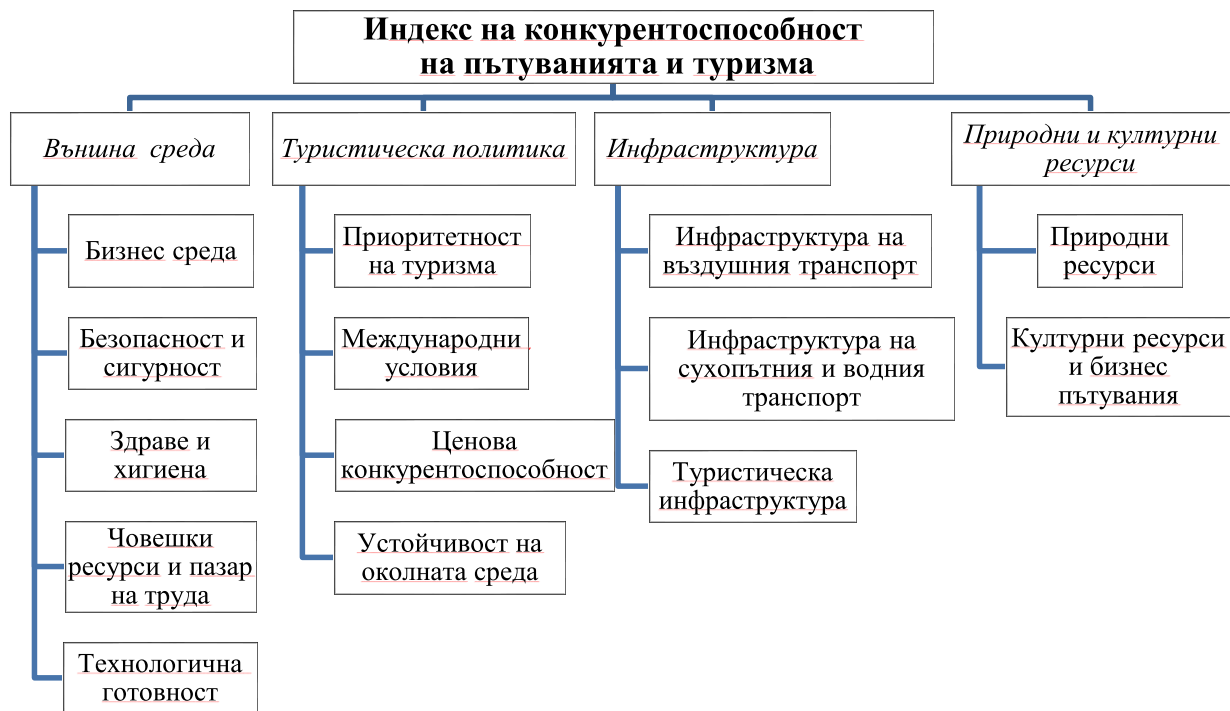
²⁶ Op. cit., The Travel... <http://www3.weforum.org>.

- „Природни и културни ресурси“.

Субиндексът „Външна среда“ предоставя информация за предпоставките, чието осигуряване е необходимо условие за развитието на туризма в отделните държави. Политиките и стратегиите, които оказват пряко въздействие върху сектора, се установяват чрез **субиндекса „Туристическа политика“**. Конкументните стълбове, чрез които се определят достъпността и качеството на физическата инфраструктура, са включени в **субиндекса „Инфраструктура“**, а тези фактори, които разкриват основните причини за пътуване до дадена страна - в **субиндекса „Природни и културни ресурси“**.

Отделните субиндекси са изградени на базата на четиринадесет стълба (критерия), чрез които се оценяват различни аспекти на глобалната конкурентоспособност на туристическата дестинация. Те са: Бизнес среда, Безопасност и сигурност, Здраве и хигиена, Човешки ресурси и пазар на труда, Технологична готовност, Приоритетност на туризма, Международни условия, Ценова конкурентоспособност, Устойчивост на околната среда, Инфраструктура на въздушния транспорт, Инфраструктура на сухопътния и водния транспорт, Туристическа инфраструктура, Природни ресурси, Културни ресурси и бизнес пътувания. Съставните компоненти на индекса на конкурентоспособността на пътуванията и туризма са представени на фиг. 1.

Фиг. 1. Базови показатели за формиране на индекса на конкурентоспособност на пътуванията и туризма



Четиринадесетте стълба са съставени от деветдесет индивидуални показателя, чиито оценки са нормирани в границите от 1 до 7, като 1 е най-ниската стойност, а 7 -

най-високата. В зависимост от посоката на влияние на отделния фактор - стимулиращо или ограничаващо, оценката се изчислява по следните формули²⁷:

а) при стимулиращо влияние:

$$V_i = 6 \times \left(\frac{X_i - X_i \min}{X_i \max - X_i \min} \right) + 1, \quad (1)$$

където:

V_i е нормираната оценка;

X_i - фактичката оценка на i -тия показател;

$X_i \min$ - минимална стойност на i -тия показател;

$X_i \max$ - максимална стойност на i -тия показател.

б) при ограничаващо влияние:

$$V_i = (-6) \times \left(\frac{X_i - X_i \min}{X_i \max - X_i \min} \right) + 7. \quad (2)$$

Оценките на елементите на всяко следващо ниво на агрегиране - от показатели към стълбове, субиндекси и индекс на конкурентоспособността, се изчисляват като непретеглени средни аритметични величини.

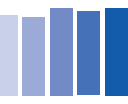
Клъстерният анализ е многомерен статистически метод, който представлява съвкупност от алгоритми, въз основа на които множество обекти могат да се групират в две или повече еднородни групи (клъстери) според значенията на определен набор от признаци. Отделените клъстери се характеризират с вътрешна хомогенност и ясно изразена външна хетерогенност.

В литературата са разработени различни варианти на клъстерния анализ. Те се класифицират в зависимост от избрания подход за определяне на броя на клъстерите (т.е. дали предварително е определен броят на клъстерите), от избраната метрика за определяне на разстоянието между отделните обекти по даден признак, както и от използваните правила за отнасянето на даден обект към конкретна група.

В настоящата статия е възприет подход, при който клъстеризацията на страните в еднородни групи е извършена на два етапа²⁸. На първия етап е проведен йерархичен клъстерен анализ, като целта е да се определи оптималният брой клъстери, в които да бъдат разпределени страните. Приложен е методът на междугруповото свързване, който е итеративна процедура по обединяване на всяка двойка най-близко разположени клъстери в нов. Разстоянието между клъстерите А и В се изчислява като

²⁷ Op. cit., The Travel... p. 30. <http://www3.weforum.org>.

²⁸ Петков, П. Клъстерен анализ при макросегментиране на международните пазари. Количествени методи в икономиката - класика и новаторство, Свищов, 2000, с. 144 - 151.



средно разстояние между всички двойки обекти в тях - n_A обекта от А и n_B обекта от В²⁹:

$$D(A, B) = \frac{1}{n_A \cdot n_B} \sum_{i=1}^{n_A} \sum_{j=1}^{n_B} d(x_i, x_j), \quad (3)$$

където $d(x_i, x_j)$ е избраната метрика за измерване на разстоянието между отделните обекти по зададените променливи.

За установяване на дистанцията между обектите е използвано евклидовото разстояние, което се изчислява по следната формула:

$$d(x_i, x_j) = \sqrt{\sum_{m=1}^k (x_{im} - x_{jm})^2}, \quad (4)$$

където:

x_{im} е значението на i -тия обект по m -тата променлива;

x_{jm} - значение на j -тия обект по m -тата променлива;

k - брой променливи, по които се извършва клъстеризацията.

Определянето на **броя на клъстерите** е от съществено значение за резултатите от анализа. В научната литература са изведени няколко подхода за изследване на валидността на клъстерите, които според **Theodoridis** и **Koutroubas**³⁰ могат да се обобщят в три групи. Първият подход се основава на външни критерии като сравняване на резултатите с такива от аналогични изследвания, субективни съображения, изследователски опит. Вторият подход се основава на конкретни резултати, получени в процеса на анализа, например разстоянието, при което се обединяват двойките клъстери. При третия подход решението се взема на базата на относителни критерии, чрез които се сравняват и оценяват структури при различен брой клъстери.

В настоящия анализ е възприет третият подход, като за определяне на оптималния брой клъстери са приложени следните индекси³¹: Calinski и Harabasz (**CH index**), Krzanowski и Lai (**KL index**), Hartigan (**Hartigan index**), Rubin (**Rubin index**) и Dunn (**Dunn index**). Тествани са варианти с два до осем на брой клъстера. Емпиричните данни за анализа през изследвания период са представени в табл. 1 и табл. 2:

²⁹ Rencher, A. C. *Methods of Multivariate Analysis*, 2nd ed., Wiley, New York, 2002, p. 463.

³⁰ Theodoridis S, Koutroubas, K. *Pattern Recognition*. 4th edition. Academic Press, 2008.

³¹ Charrad, M., Ghazzali, N., Boiteau, V., Niknafs, A. NbClust: An R Package for determining the relevant number of clusters in a data set. // *Journal of Statistical Software*, 2014, Vol. 61 (6), pp. 5 - 18.

1. Оценки на стълбовете на конкурентоспособността през 2015 година

Страни	BI	SS	HH	HRLB	ICT	PTT	IO	PC	ES	ATI	GPI	TSI	NR	CRB
Австрия	4.94	6.47	6.97	5.09	5.70	5.52	3.99	3.49	5.13	4.01	5.42	6.83	4.07	2.92
Албания	4.11	5.34	5.22	4.68	4.07	4.03	2.34	4.38	3.60	2.16	3.01	3.94	2.03	1.14
Белгия	4.71	6.18	6.49	5.03	5.47	4.47	4.13	3.73	4.28	3.83	5.96	5.65	2.65	3.67
Бивша югославска република Македония	4.87	5.75	5.99	4.47	4.47	4.41	2.36	4.55	3.65	2.39	3.25	4.58	2.15	1.30
България	4.22	5.24	6.70	4.72	4.76	4.18	3.87	5.08	4.62	2.46	3.26	6.06	3.44	1.96
Германия	6.32	6.06	6.85	5.18	5.51	4.84	4.24	3.62	4.90	4.93	5.99	5.61	4.41	6.00
Гърция	4.04	5.49	6.57	4.75	4.71	5.36	4.09	3.93	4.19	4.25	4.01	6.08	3.49	2.82
Дания	5.28	5.88	6.11	5.47	6.18	4.31	4.34	3.31	4.92	3.87	5.52	4.67	3.45	2.18
Естония	5.13	6.04	6.25	5.12	5.71	5.76	3.65	4.62	4.73	2.97	4.39	5.87	2.69	1.55
Ирландия	5.37	6.18	5.80	5.27	5.28	5.25	4.53	3.69	5.31	4.16	4.89	6.10	2.79	2.82
Исландия	4.96	6.54	6.07	5.49	5.88	5.89	4.38	3.59	4.92	4.67	4.35	6.28	3.63	1.53
Испания	4.09	5.97	6.11	4.87	5.26	5.89	3.93	4.22	4.61	4.91	5.54	6.58	4.59	6.69
Италия	3.59	5.68	6.27	4.45	5.14	4.62	4.09	3.49	4.34	4.26	4.65	6.66	4.60	6.51
Кипър	4.72	6.00	5.80	5.16	4.63	5.96	3.75	3.97	3.92	3.48	4.91	6.77	2.66	1.88
Латвия	4.59	5.79	6.17	5.18	5.60	4.64	3.97	4.84	4.59	3.12	4.22	5.08	2.55	1.33
Литва	4.48	5.56	6.81	4.96	5.29	4.37	3.99	4.87	4.22	2.39	4.55	4.24	2.44	1.50
Люксембург	5.73	6.46	6.26	5.16	6.09	4.64	4.25	4.10	5.23	3.59	5.73	4.67	3.09	1.62
Малта	4.76	6.03	6.35	4.61	5.24	6.03	3.94	4.22	4.17	3.83	4.98	5.61	2.22	1.44
Молдова	3.90	5.43	6.26	4.50	4.23	3.82	1.96	4.80	4.22	1.90	2.77	3.29	1.75	1.08
Нидерландия	5.44	6.16	6.24	5.13	5.96	4.66	4.25	3.56	4.81	4.89	6.21	4.62	3.19	3.51
Норвегия	5.44	6.10	6.17	5.24	6.14	5.14	3.97	3.23	5.22	5.01	3.81	5.49	3.96	2.22
Обединено кралство	5.70	5.44	5.83	5.29	6.09	5.10	4.24	2.73	4.79	5.12	5.51	5.08	4.79	5.90
Полша	4.35	5.86	6.21	4.80	4.90	4.10	4.08	4.94	4.62	2.57	4.08	4.44	3.14	2.77
Португалия	4.54	6.33	6.06	5.18	4.97	5.46	4.21	4.23	4.42	3.91	4.54	6.12	3.70	3.71
Румъния	4.11	5.42	5.94	4.56	4.36	4.34	3.91	4.89	4.35	2.34	3.10	5.01	2.70	2.07
Руска федерация	3.98	3.95	6.69	4.83	4.83	4.33	2.48	4.99	3.70	4.42	3.09	4.65	3.77	3.32
Словакия	3.92	5.55	6.42	4.75	5.05	4.04	3.89	4.51	4.49	1.78	4.22	4.94	3.31	1.42
Словения	4.03	6.20	6.05	4.69	5.07	4.93	3.70	4.34	4.74	2.44	5.13	5.72	3.87	1.40
Сърбия	3.38	5.46	6.04	4.29	4.45	3.83	2.39	4.56	4.08	2.13	2.95	4.50	1.90	1.61
Турция	4.54	4.23	5.41	4.30	4.17	4.39	3.34	4.37	3.83	4.66	3.88	5.04	2.78	3.83
Унгария	4.28	5.79	6.61	4.79	4.93	5.13	4.15	4.60	5.16	2.71	4.45	5.02	2.72	2.22
Финландия	5.60	6.70	6.31	5.43	6.37	4.57	4.10	3.71	5.25	4.41	4.83	5.02	3.16	2.13
Франция	4.52	5.44	6.52	4.96	5.55	5.16	4.22	2.95	4.67	4.98	5.78	6.15	4.80	6.56
Хърватия	3.65	6.00	6.33	4.41	5.03	4.51	4.14	4.28	4.38	3.06	4.20	6.35	3.80	2.67
Черна гора	4.39	5.69	5.71	4.85	4.42	4.57	2.44	4.48	4.30	3.03	3.51	5.84	2.76	1.09
Чешка република	4.35	5.71	6.73	4.75	5.19	4.61	4.15	4.47	4.90	3.13	5.15	5.44	2.59	2.30
Швейцария	5.76	6.32	6.50	5.64	6.03	5.64	4.11	2.57	5.63	5.03	6.00	6.35	4.32	2.93
Швеция	5.22	6.10	5.94	5.30	6.17	4.70	4.07	3.38	5.03	4.52	4.76	4.84	3.20	2.93

Източник: The Travel & Tourism Competitiveness Report 2015.

Използвани съкращения:

Бизнес среда (BI), Безопасност и сигурност (SS), Здраве и хигиена (HH), Човешки ресурси и пазар на труда (HRLB), Технологична готовност (ICT), Приоритетност на туризма (PTT), Международни условия (IO), Ценова конкурентоспособност (PC), Устойчивост на околната среда (ES), Инфраструктура на въздушния транспорт (ATI), Инфраструктура на сухопътния и водния транспорт (GPI), Туристическа инфраструктура (TSI), Природни ресурси (NR), Културни ресурси и бизнес пътувания (CRB).

2. Оценки на стълбовете на конкурентоспособността през 2017 година

Страни	BI	SS	HH	HRLB	ICT	PTT	IO	PC	ES	ATI	GPI	TSI	NR	CRB
Австрия	4.96	6.28	6.71	5.52	5.84	5.33	4.01	3.86	5.63	3.89	5.25	6.67	4.11	3.06
Албания	4.07	5.71	5.22	4.90	4.14	4.56	2.40	4.70	4.11	2.00	3.06	3.91	2.19	1.14
Белгия	4.93	5.90	6.71	5.35	5.72	4.47	4.15	4.38	4.57	3.74	5.70	5.29	2.20	4.09
Бивша югославска република Македония	4.82	5.57	5.99	4.42	4.63	4.34	2.64	5.20	3.74	2.18	3.30	3.99	2.14	1.35
България	4.50	5.14	6.63	4.73	5.03	4.27	3.89	5.31	4.97	2.37	3.11	5.80	3.82	2.12
Германия	5.34	5.65	6.86	5.55	5.85	4.76	4.26	4.17	5.19	4.92	5.76	6.00	3.98	6.28
Гърция	4.11	5.63	6.57	4.80	4.92	5.47	4.11	4.66	4.48	4.27	3.75	5.70	4.06	3.08
Дания	5.48	6.05	6.13	5.66	6.41	4.44	4.36	3.75	5.17	3.49	5.44	4.77	3.27	2.28
Естония	5.23	6.26	6.27	5.20	6.10	5.45	3.67	5.07	4.90	3.01	4.45	5.46	2.45	1.59
Ирландия	5.51	6.11	5.71	5.45	5.65	5.41	4.51	4.03	4.70	4.16	4.70	5.76	2.84	2.91
Исландия	5.33	6.57	6.07	5.76	6.11	6.03	4.40	3.58	4.85	4.69	3.97	5.82	3.50	1.47
Испания	4.39	6.16	6.28	4.94	5.50	5.91	3.95	4.52	4.63	5.00	5.19	6.66	4.91	6.85
Италия	3.87	5.41	6.18	4.61	5.37	4.54	4.13	3.88	4.49	4.36	4.67	5.96	4.78	6.46
Кипър	4.59	5.82	5.81	4.92	4.77	5.71	3.77	4.27	4.05	3.05	3.67	5.63	2.90	1.78
Латвия	4.56	5.76	6.44	5.01	5.31	4.47	3.99	5.25	4.88	3.05	4.05	4.64	2.38	1.41
Литва	4.63	5.73	6.81	5.02	5.46	4.27	4.01	5.37	4.45	2.40	4.39	4.42	2.23	1.50
Люксембург	5.80	6.32	6.27	5.35	6.16	4.84	4.27	4.55	5.48	3.58	5.53	5.91	2.65	1.74
Малта	4.75	5.92	6.35	4.76	5.43	6.18	3.96	4.44	4.09	3.86	4.53	5.50	3.08	1.49
Молдова	3.81	5.42	6.13	4.31	4.30	3.44	2.13	5.38	4.06	1.98	2.55	2.80	1.60	1.16
Нидерландия	5.50	6.14	6.23	5.52	6.09	4.71	4.27	4.13	5.10	4.95	6.13	4.86	2.25	3.36
Норвегия	5.56	6.41	6.31	5.63	6.27	5.21	4.05	3.69	5.55	5.28	3.48	5.43	4.10	2.22
Обединено кралство	5.86	5.34	5.83	5.50	6.19	4.98	4.24	2.83	4.70	5.20	5.36	6.16	4.64	5.96
Полша	4.47	5.70	6.25	4.87	5.06	4.15	4.10	5.50	4.58	2.61	4.27	4.19	2.98	2.84
Португалия	4.63	6.32	6.25	5.16	5.22	5.53	4.23	4.83	4.35	3.94	4.19	6.37	3.88	3.89
Румъния	4.36	5.78	6.14	4.45	4.70	3.81	3.93	4.68	4.41	2.37	2.80	4.37	2.99	2.27
Руска федерация	4.09	4.31	6.70	4.83	4.98	4.21	2.21	5.75	4.11	4.51	3.05	4.46	3.85	3.23
Словакия	4.01	5.61	6.50	4.70	5.37	4.08	3.91	4.95	4.84	1.75	4.19	4.34	3.43	1.53
Словения	4.33	6.20	6.03	4.91	5.22	4.83	3.72	4.63	5.07	2.46	4.76	5.40	3.78	1.46
Сърбия	4.02	5.41	6.04	4.43	4.80	3.60	2.41	4.82	4.18	2.35	2.77	3.92	2.01	1.65
Турция	4.49	4.05	5.43	4.26	4.31	4.32	3.86	4.86	3.68	4.74	3.54	4.74	2.98	4.10
Унгария	4.17	5.72	6.63	4.66	4.88	4.90	4.17	4.66	4.71	3.00	4.45	4.44	2.61	2.27
Финландия	5.65	6.65	6.29	5.55	6.19	4.59	4.12	4.26	5.45	3.99	4.64	4.68	2.88	2.09
Франция	4.74	5.43	6.52	5.14	5.86	5.14	4.24	4.12	4.83	4.90	5.65	5.70	4.77	6.75
Хърватия	4.01	6.09	6.38	4.39	5.05	4.50	4.16	4.45	4.73	3.05	3.89	6.26	4.50	2.77
Черна гора	4.36	5.37	5.80	4.46	4.83	4.64	2.44	4.82	4.33	3.01	3.23	5.36	2.58	1.11
Чешка република	4.46	5.92	6.69	5.00	5.60	4.21	4.17	4.90	4.87	3.10	4.94	5.10	2.47	2.41
Швейцария	6.00	6.41	6.53	5.74	6.35	5.58	4.13	2.81	5.80	4.94	5.89	6.20	3.74	2.87
Швеция	5.46	6.22	6.11	5.53	6.34	4.62	4.09	4.05	5.32	4.59	4.59	5.04	3.08	2.97

Източник: The Travel & Tourism Competitiveness Report 2017.

Оптималният брой кълстери според индексите CH, KL и Dunn съответства на максималните им стойности, а според индексите на Hartigan - на максималната разлика между поредните индекси, и Rubin - на минималната стойност на вторите разлики между индексите. Емпиричните стойности и на петте индекса дават съгласувано заключение относно обособяването на пет кълстера през 2015 г. и на три през 2017 г. като най-подходящо решение (табл. 3).

3. Емпирични значения на индексите за определяне на оптимален брой кълстери

Брой кълстери	CH index		KL index		Hartigan index		Rubin index		Dunn index	
	2015	2017	2015	2017	2015	2017	2015	2017	2015	2017
2	16.67	17.59	1.49	1.34	12.74	15.51	51.70	51.68	0.39	0.26
3	17.18	19.77	3.17	2.02	4.97	9.10	70.01	73.95	0.35	0.29
4	14.31	19.08	0.36	1.71	14.05	5.99	79.94	93.18	0.35	0.26
5	18.13	17.79	7.35	1.21	2.66	5.26	112.98	109.58	0.51	0.28
6	15.72	17.02	0.40	1.28	5.66	4.41	122.09	127.03	0.50	0.26
7	15.85	16.34	1.44	1.74	4.29	2.79	143.68	144.53	0.45	0.26
8	15.56	15.16	1.22	0.96	3.75	2.88	163.61	157.52	0.48	0.25

Източник: собствени изчисления.

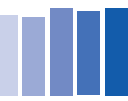
На втория етап от анализа чрез метода на k-средните е извършено разпределението на европейските държави в съответния брой кълстери за двете изследвани години. Направена е проверка чрез дисперсионен анализ относно значимостта на признаците, въз основа на които е извършена кълстеризацията. Резултатите показват, че влиянието на всички конкурентни стълбове е статистически значимо с изключение на „Здраве и хигиена“ (табл. 4).

4. Значимост на факторите за различията между кълстерите

Променливи	2015		2017		Променливи	2015		2017	
	F	Sig.	F	Sig.		F	Sig.	F	Sig.
Бизнес среда	7.01	0.00	10.55	0.00	Ценова конкурентоспособност	18.47	0.00	14.31	0.00
Безопасност и сигурност	32.76	0.00	13.31	0.00	Устойчивост на околната среда	14.53	0.00	8.05	0.00
Здраве и хигиена	2.30	0.08	0.99	0.38	Инфраструктура на въздушния транспорт	25.55	0.00	21.93	0.00
Човешки ресурси и пазар на труда	11.90	0.00	15.04	0.00	Инфраструктура на сухоземния и водния транспорт	18.76	0.00	17.03	0.00
Технологична готовност	30.88	0.00	24.28	0.00	Туристическа инфраструктура	3.28	0.02	16.46	0.00
Туристическа политика	3.09	0.03	7.99	0.00	Природни ресурси	16.81	0.00	14.01	0.00
Международни условия	31.76	0.00	10.47	0.00	Културни ресурси и бизнес пътувания	52.89	0.00	65.65	0.00

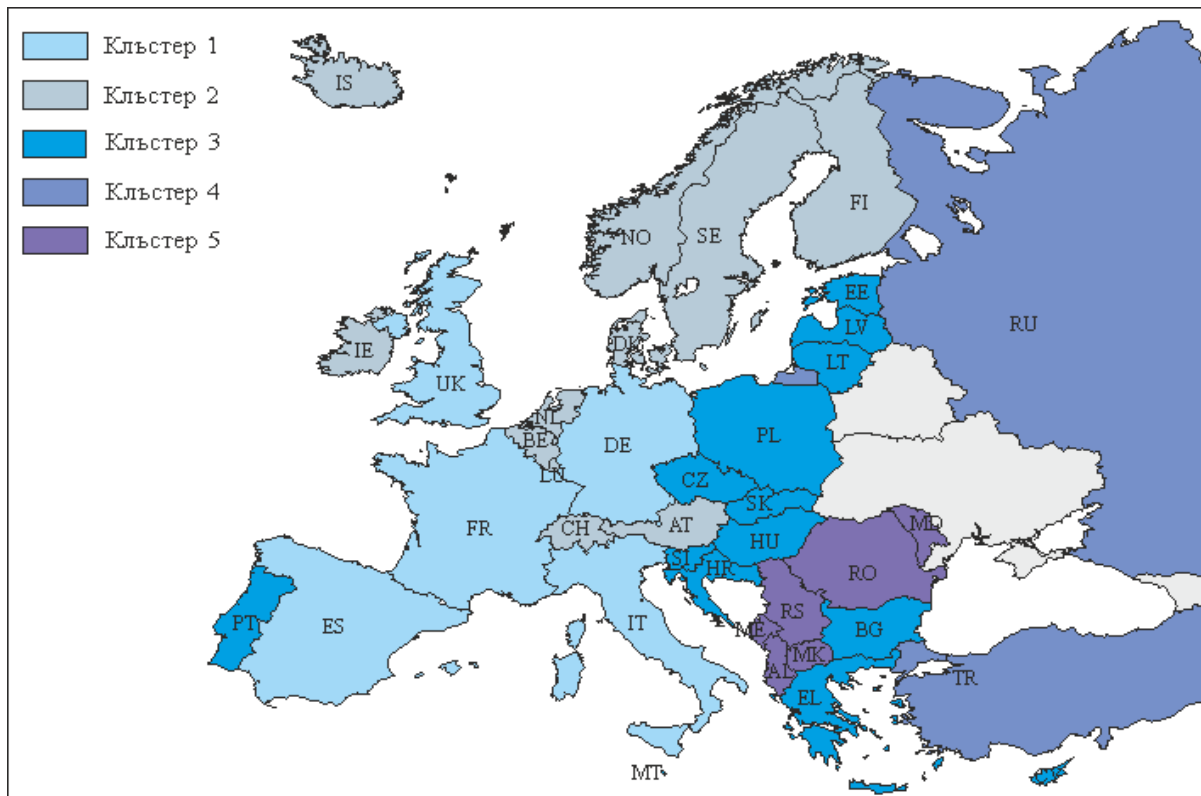
Източник: собствени изчисления.

От проведения кълстерен анализ се установи, че през 2015 г. се обособяват следните пет еднородни групи страни: с много висока конкурентоспособност (*кълстер 1*), с висока конкурентоспособност (*кълстер 2*), със средно равнище на конкурентоспособност (*кълстер 3*), с ниска конкурентоспособност (*кълстер 4*) и с



много ниска конкурентоспособност (клъстер 5). Разпределението на включените в анализа европейски държави в така обособените клъстери е представено на фиг. 2.

Фиг. 2. Разпределение на европейските държави по клъстери според индекса на конкурентоспособност на пътуванията и туризма през 2015 година



Източник: авторът.

- **Първи клъстер** - Испания, Франция, Германия, Италия, Обединено кралство. В този клъстер са включени страни с най-високи стойности на показателите, които формират субиндексите „Инфраструктура“ и „Природни и културни ресурси“, както и високи оценки по останалите критерии за клъстеризация. Като единствена слаба страна на държавите от клъстера се определя ниското равнище на ценовата конкурентоспособност.

- **Втори клъстер** - Швейцария, Австрия, Нидерландия, Белгия, Люксембург, Исландия, Ирландия, Норвегия, Финландия, Швеция, Дания. Държавите в клъстера се характеризират с благоприятна среда за развитие на туристическия сектор, както и с висока степен на международна интегрираност и устойчивост на околната среда. Като слаба страна на държавите от този клъстер, както и при клъстер 1, се откроява ниската ценова конкурентоспособност.

- **Трети клъстер** - Португалия, Гърция, Хърватия, Кипър, Словения, Малта, Чешка република, Естония, Унгария, Полша, България, Латвия, Литва, Словакия. В този клъстер са включени предимно страни от Централна и Източна Европа, които се характеризират с недостатъчно добре развита и функционираща инфраструктура и подкрепяща развитието на туризма среда. Значително е изоставането на държавите

спрямо страните от първия клъстер по отношение на конкурентните стълбове, включени в субиндекса „Културни ресурси и бизнес пътувания“.

- **Четвърти клъстер** - Руска федерация, Турция. В този клъстер са включени държави, чиито силни страни се отнасят единствено до критериите „Ценова конкурентоспособност“, „Инфраструктура на въздушния транспорт“ и „Културни ресурси и бизнес пътувания“. Като най-слаба страна се отчита безопасността и сигурността, чието равнище е най-ниско сред изследваните държави.

- **Пети клъстер** - Черна гора, бивша югославска република Македония, Сърбия, Албания, Румъния, Република Молдова. Страните от този клъстер изостават значително от останалите европейски държави по отношение на всички стълбове. Единственото конкурентно предимство на страните от клъстера се отнася до цените на туристическите услуги.

Евклидовите разстояния (d) между така обособените клъстери (табл. 5) показват, че най-малка е дистанцията между клъстер 2 и клъстер 3, а най-голяма - между клъстер 1 и клъстер 5. Първият клъстер в най-голяма степен се различава от останалите, тъй като средното разстояние до всеки от останалите клъстери, измерено отново чрез евклидовото разстояние, е най-голямо. По отношение на дистанцията между отделните страни и съответстващите им клъстерни центрове бихме могли да определим като най-еднороден клъстер 4, а като най-разнороден - клъстер 3.

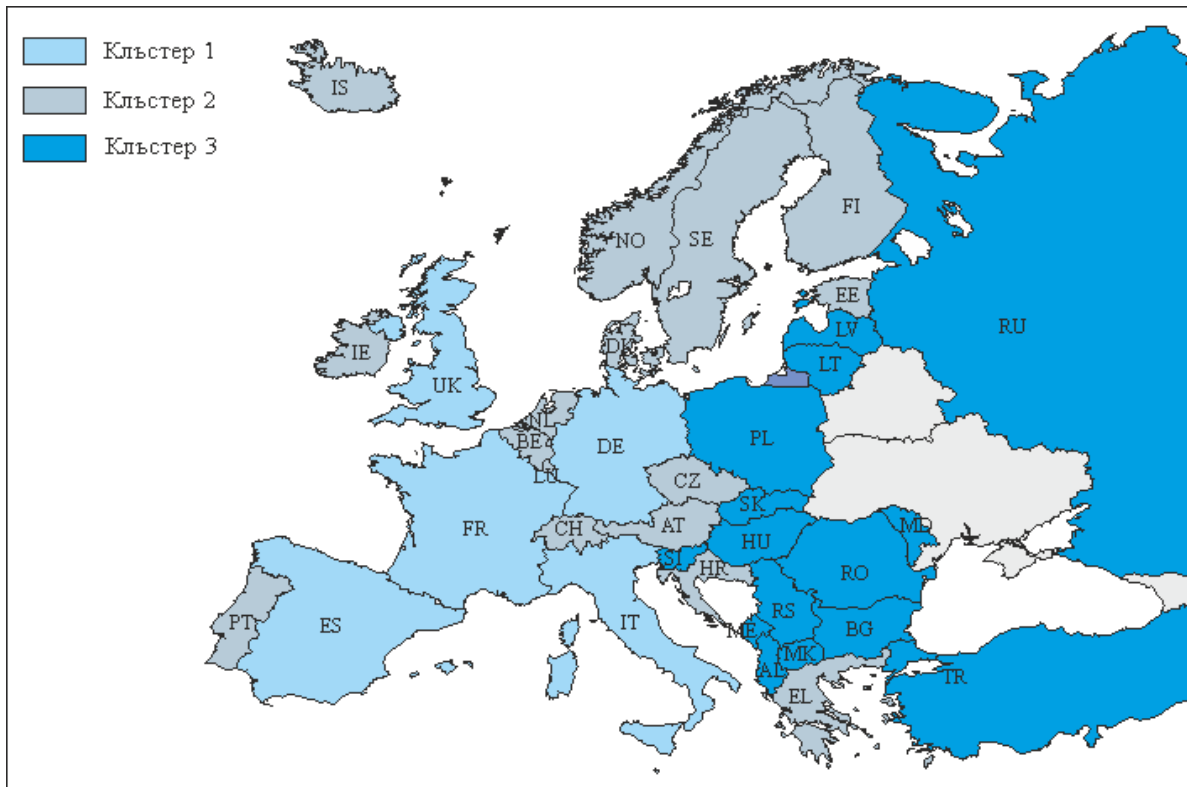
5. Евклидово разстояние между клъстерите

Клъстери	d_{2015}	Клъстери	d_{2015}	Клъстери	d_{2017}
$K_1 - K_2$	4.14	$K_2 - K_4$	4.26	$K_1 - K_2$	4.32
$K_1 - K_3$	5.21	$K_2 - K_5$	4.84	$K_1 - K_3$	6.15
$K_1 - K_4$	4.87	$K_3 - K_4$	3.41	$K_2 - K_3$	3.12
$K_1 - K_5$	7.23	$K_3 - K_5$	2.88	-	-
$K_2 - K_3$	2.51	$K_4 - K_5$	3.67	-	-

Източник: собствени изчисления.

През 2017 г. се наблюдава намаляване на различията между част от включените в анализа държави, в резултат на което се формират три клъстера (фиг. 3). Съставът на клъстер 1 се запазва без промени, като същевременно с това се отчита повишаване на хомогенността на групата, индикация за което е по-ниската стойност на средното вътреклъстерно разстояние. По-ясно изразени промени се наблюдават в клъстер 2, към състава на който се присъединяват шест страни от клъстер 3: Португалия, Малта, Чешката република, Естония, Гърция и Хърватия. В резултат се отчита минимално раздалечаване на страните в клъстера, тъй като средното разстояние до клъстерния център се увеличава с 11%.

Фиг. 3. Разпределение на европейските държави по клъстери според индекса на конкурентоспособност на пътуванията и туризма през 2015 година



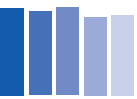
Източник: авторът.

Най-голямо изменение се наблюдава в клъстер 3, където преминават всички страни от обособените през 2015 г. клъстери 4 и 5. В състава на така формирания нов клъстер влизат следните шестнадесет държави: Кипър, Словения, Унгария, Полша, България, Латвия, Литва, Словакия, Руска федерация, Турция, Черна гора, бившата югославска република Македония, Сърбия, Албания, Румъния и Република Молдова. Този клъстер остава най-хетерогенен, като след включването в неговия състав на страните от несъществуващите вече два клъстера се отчита и увеличение с 13.6% на дистанцията между отделните страни и съответстващите им клъстерни центрове.

Заклучение

Анализът на конкурентоспособността на туристическия сектор в отделните европейски държави въз основа на стълбовете на индекса на конкурентоспособност на пътуванията и туризма показва, че между страните съществуват сериозни различия в степента на развитие на сектора. В резултат от клъстеризацията се установи, че страните с най-високи индекси на конкурентоспособност са тези, които освен че притежават културно-исторически ценности, в максимална степен отговарят на съвременните тенденции за развитието на туристическия бизнес.

За успешното развитие на туристическия сектор в България е необходимо наличните сравнителни предимства да се трансформират в конкурентни, които към момента основно се отнасят до ниските цени на туристическите услуги, природните ресурси и богатото културно-историческо наследство. Същевременно е необходимо да



се работи в посока подобряване на бизнес средата и качеството на предоставяните услуги, както и за развитие на инфраструктурата. Изпълнението на тези условия би довело до конкурентоспособен и ефективен туристически сектор в страната, чиито ползи за населението имат не само икономически, но и социален характер.

ЦИТИРАНА ЛИТЕРАТУРА:

- Иванов, Л.** (2008). Статистическо изследване и прогнозиране на развитието. Свищов.
- Мандель, И. Д.** (1988). Кластерный анализ. Финансы и статистика.
- Петков, П.** (2000). Клъстерен анализ при макросегментиране на международните пазари. Количествени методи в икономиката - класика и новаторство, Свищов.
- Хаджиев, В.,** В. Димитрова, Л. Любенов (2009). Статистически и иконометричен софтуер, 3-то пр. и доп. изд. Наука и икономика, Варна.
- Charrad, M.,** N. Ghazzali, V. Boiteau, A. Niknafs (2014). NbClust: An R Package for determining the relevant number of clusters in a data set. // *Journal of Statistical Software*, Vol. 61 (6).
- Chon K. S.,** K. J. Mayer (1995). Destination competitiveness models in tourism and their application to Las Vegas. // *Journal of Tourism Systems and Quality Management*, Vol. 1, No. 2/4.
- Crouch, G. I.** (2007). Measuring tourism competitiveness: research, theory and the WEF index. Proceedings of the ANZMAC.
- Crouch, G. I.** (2011). Destination Competitiveness: An Analysis of Determinant Attributes. // *Journal of Travel Research*, Vol. 50 (1).
- D’Hautesserre A. M.** (2000). Lessons in managed destination competitiveness: The case of Foxwoods Casino Resort. // *Tourism Management*, Vol. 21 (1).
- De Holan, P. M.,** N. Phillips (1997). Sun, sand, and hard currency: Tourism in Cuba. // *Annals of Tourism Research*, Vol. 24 (4), pp. 777 - 795.
- De Keyser, R.,** N. Vanhove (1994). The competitive situation of tourism in the Caribbean area methodological approach. // *Revue de Tourisme*, Vol. 49 (3).
- Dupeyras, A.,** N. MacCallum, Indicators for Measuring Competitiveness in Tourism: A Guidance Document, OECD Tourism Papers, 2013/02, OECD Publishing.
- Dwyer, L. P.,** P. Forsyth, P. Rao (2000). The price competitiveness of travel and tourism: a comparison of 19 destinations. // *Tourism Management*, No. 21.
- Dwyer, L.,** C. Kim (2003). Destination competitiveness: determinants and indicators. // *Current issues in tourism*, Vol. 6 (5).
- Enright, M.,** J. Newtown (2005). Determinants of tourism destination competitiveness in Asia Pacific: Comprehensiveness and universality. // *Journal of Travel Research*, No. 43.
- Everitt, B.,** S. Landau, M. Leese, D. Stahl (2011). Cluster analysis, fifth edition. John Wiley & Sons, Inc.
- Hassan, S.** (2000). Determinants of market competitiveness in an environmentally sustainable tourism industry. // *Journal of Travel Research*, No. 38.
- Heath, E.** (2002). Towards a model to enhance destination competitiveness: A Southern African perspective. // *Journal of Hospitality and Tourism Management*, Vol. 10 (2).
- Hong, S.** (2008). Competitiveness in the Tourism Sector, A Comprehensive Approach from Economic and Management Points. Springer Science and Business Media.

Mazanec, J. A., A. Ring (2011). Tourism destination competitiveness: second thoughts on the World Economic Forum reports. // *Tourism Economics*, Vol. 17 (4).

Mihalic, T. (2000). Environmental management of a tourist destination A factor of tourism competitiveness. // *Tourism Management*, No. 21.

Navickas, V., A. Malakauskaite (2009). The Possibilities for the Identification and Evaluation of Tourism Sector Competitiveness Factors. // *Engineering Economics*, Vol. 61 (1).

Poon, A. (1993). Tourism, technology and competitive strategies. CAB international.

Rencher, A. C. (2002). Methods of Multivariate Analysis, 2nd ed., Wiley, New York.

Ritchie J. R. B., G. I. Crouch (1999). Tourism, competitiveness, and societal prosperity. // *Journal of Business Research*, Vol. 44 (3).

Ritchie J. R. B., G. I. Crouch (2000). The competitive destination: A sustainability perspective. // *Tourism Management*, Vol. 21 (1).

Ritchie J. R. B., G. I. Crouch (2003). The competitive destination: A sustainable tourism perspective, CAB International, Wallingford.

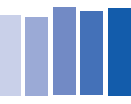
Theodoridis S., K. Koutroubas (2008). Pattern Recognition. 4th edition. Academic Press.

The Travel and Tourism Competitiveness Report. <http://www3.weforum.org>.

Tourism towards 2030. Global overview. <http://www2.unwto.org/>.

Travel and Tourism Economic Impact 2017 World. <https://www.wttc.org/>.

Yoon, Y. (2000). Development of a structural model for tourism destination competitiveness from stakeholders' perspectives (Doctoral dissertation, Virginia Tech).



КЛЪСТЕРИЗАЦИЯ НА ЕВРОПЕЙСКИТЕ СТРАНИ СПОРЕД СТЕПЕНТА НА ТУРИСТИЧЕСКА КОНКУРЕНТОСПОСОБНОСТ

*Тихомир Върбанов**

РЕЗЮМЕ В статията се изследва конкурентоспособността на европейските страни в сферата на туризма, оценена чрез система от показатели, формиращи индекса за конкурентоспособност на пътуванията и туризма. Използван е клъстерен анализ за обособяване на еднородни групи страни и са идентифицирани факторите с най-голямо влияние при тяхното формиране. Установено е мястото на България и са очертани необходимите действия за повишаване на конкурентоспособността и устойчивото развитие на страната като туристическа дестинация.

Ключови думи: индекс на конкурентоспособност на пътуванията и туризма, клъстерен анализ, критерии, индикатори.

JEL: C10, L83

* Асистент д-р в катедра „Математика и статистика“, СА „Д. А. Ценов“ - Свищов; e-mail: t.varbanov@uni-svishtov.bg.

КЛАСТЕРИЗАЦИЯ ЕВРОПЕЙСКИХ СТРАН ПО СТЕПЕНИ КОНКУРЕНТОСПОСОБНОСТИ В СФЕРЕ ТУРИЗМА

Тихомир Вырбанов

РЕЗЮМЕ В статье исследуется конкурентоспособность европейских стран в сфере туризма, с использованием системы показателей, формирующих индекс конкурентоспособности путешествий и туризма. Для определения однородных групп стран и выявления факторов, оказывающих наибольшее влияние на их формирование, был использован кластерный анализ. Было определено какое место занимает Болгария, а также предложены меры по повышению конкурентоспособности и устойчивого развития страны как туристической дестинации.

Ключевые слова: индекс конкурентоспособности путешествий и туризма, кластерный анализ, критерии, индикаторы.

JEL: C10, L83

* Доктор, асистент кафедры „Математика и статистика“, СА „Д. А. Ценов“ - Свищов; e-mail: t.varbanov@uni-svistov.bg



CLUSTERIZATION OF THE EUROPEAN COUNTRIES BASED ON THE LEVEL OF TOURISM COMPETITIVENESS

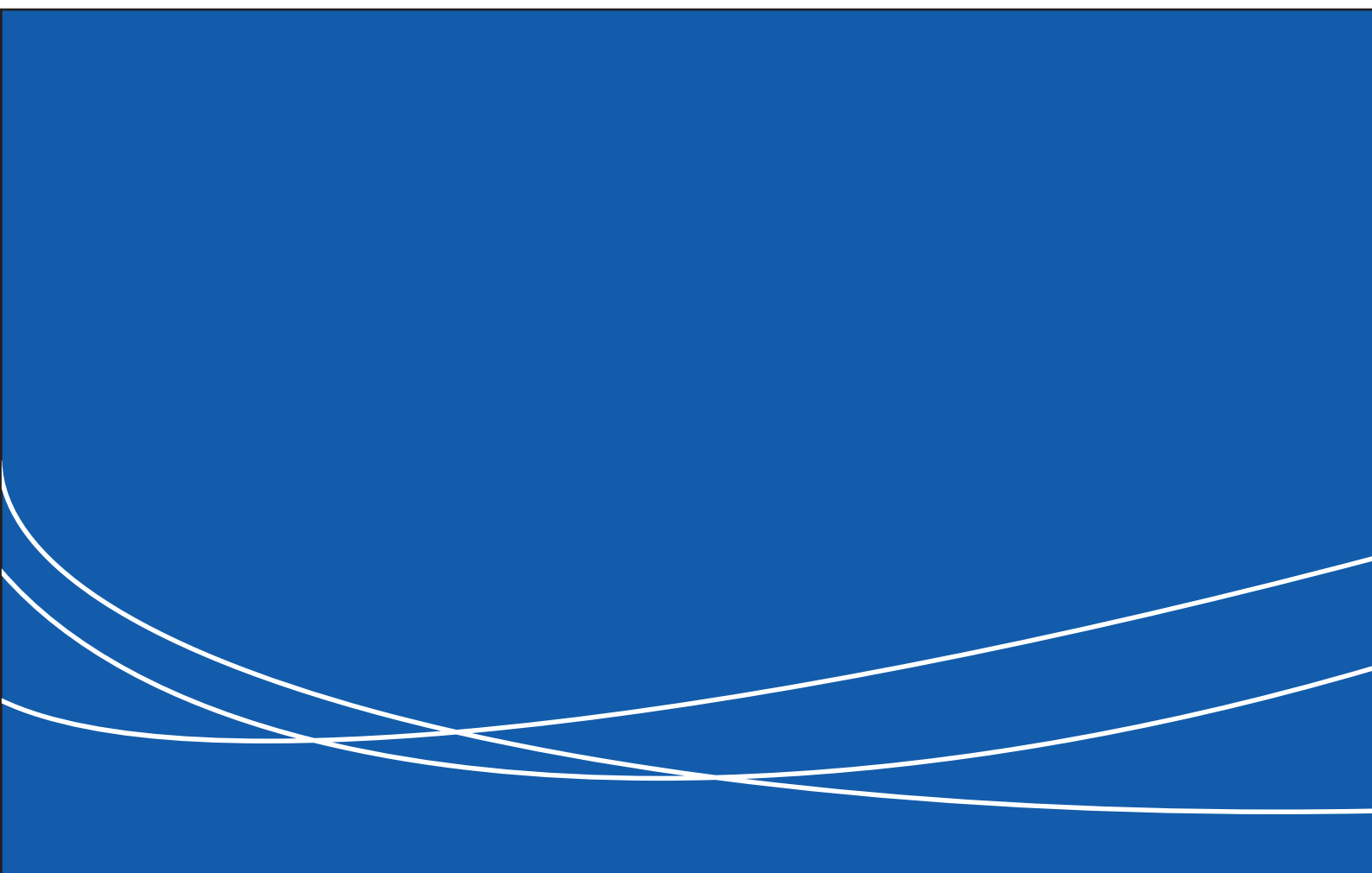
*Tihomir Varbanov**

SUMMARY The article discusses tourism competitiveness of the European countries, which is measured by a system of indicators, constituting the travel and tourism competitiveness index. Cluster analysis is used to classify the European countries in homogeneous groups. The factors with the greatest influence on the clusters' formation are identified. The place of Bulgaria is determined and the necessary actions for increasing competitiveness and sustainable development of the country as a tourism destination are outlined.

Key words: Travel and Tourism Competitiveness Index, Cluster Analysis, Criteria, Indicators.
JEL: C10, L83

* Assistant Professor at the Department of Mathematics and Statistics, Academy of Sciences 'D. A. Tsenov' - Svishtov; e-mail: t.varbanov@uni-svishtov.bg.

**БЪЛГАРСКАТА СТАТИСТИКА В ТРИ
СТОЛЕТИЯ**



ПРОФ. ВЕНЕЦ ЦОНЕВ - ЕДИН ЗАБЕЛЕЖИТЕЛЕН ТВОРЧЕСКИ ПЪТ В ТЕОРИЯТА И ПРАКТИКАТА НА СЪВКУПНОСТНИТЕ ИЗСЛЕДВАНИЯ

*Тодор Къналиев**



Тази научна конференция е посветена на ролята и постиженията на двама изтъкнати представители на статистическата наука и практика у нас, дългогодишни преподаватели в катедра „Статистика“, във връзка със стогодишнината от рождението им. Това са професорите Стефан Станев и Венец Цонев. По този начин тези от нас (а те вече са малко на брой), които са били техни студенти и колеги, имали са дългогодишни човешки и професионални контакти, могат да изразят, макар и пристрастно, своите чувства на признателност, дълбоко преклонение и съпричастие към стореното от тях през дългия им житейски път. А младите, посветили се на статистическата наука и практика, които едва ли познават физически двамата професори, биха могли да проявят любопитство и интерес към тяхното широкомащабно творчество, свързано с разработването на съществени проблеми на теорията и практиката на статистическите (съвкупностните) изследвания. Редица от тези творчески разработки имат не само вътрешнонационален характер, а и международни измерения. Важно е да се отбележи дълголетният и непрестанен интерес на двамата професори към решаването на проблемите на статистическата наука, към проблемите на преподаването по статистика в специалността „Статистика и иконометрия“ и подготовката на научни и изпълнителски кадри по статистика. За пример е обстоятелството, че техният творчески професионален живот продължи почти до смъртта им - месеци след като навършиха 90 години.

* Доц. д-р, УНСС, катедра „Статистика и иконометрия“; e-mail: todor_kan@abv.bg.

А сега конкретно за проф. Венец Цонев.

Доста трудно е да се опише и характеризира достатъчно всеотдайно човешкият и професионалният образ, както и творческите постижения на моя дълбоко уважаван учител, колега и, въпреки сериозната разлика във възрастта ни, с течение на времето, приятел проф. Венец Цонев. Мога да твърдя, че познанството и сътрудничеството ни започна от студентската скамейка в средата на 60-те години и продължи до последните му дни през 2008 година. Това ми дава достатъчно пълна представа както за човешките му качества, така и за преподавателската и научно-изследователската му работа, а също и за работата му в практиката. Ще отбележа, че в един ранен период на 60-те и 70-те години в катедра „Статистика и иконометрия“ преподаваше забележителна плеяда от преподаватели - личности като акад. Иван Стефанов, проф. Анастас Ю. Тотев, проф. Стефан Станев, проф. Здравко Сугарев, проф. Румен Янакиев, проф. Никола Наумов и други. Но сред тях проф. Венец Цонев имаше определяща роля за развитието на катедрения колектив и за усъвършенстване на обучението на студентите в специалност „Статистика“. Неговите оценки и мнения, изразени както при дискусиите по различни научни проблеми в катедрения колектив, така и по проблемите на преподаването и обучението по статистическите дисциплини в различните специалности винаги се посрещаха с интерес и се вземаха под внимание при формирането на различни решения. Авторитетът му беше много висок и дълги години работеше като неформален заместник-ръководител на катедрата при решаването на важни текущи проблеми, свързани с развитието на катедрения колектив и с обучението по статистика. Неговата всеотдайност и преданост към изпълняваните задължения, проявяваната висока ерудиция при обсъждането на различни въпроси в областта на статистиката респектираха събеседниците. Той живя до последните си дни чрез множеството си публикации с проблемите на науката, практиката и преподаването по статистика и подготовката на кадри в тази област. Тук трябва да отбележа, че през неговия дълъг и доста проблематичен професионален живот не беше станал суховат учен статистик, а беше високо ерудиран човек, с богата обща култура и със заразително свежо чувство за хумор.

Проф. В. Цонев е роден на 15 октомври 1917 г. в гр. София в семейство на държавен служител. Баща му е защитил докторат по финанси в Университета в гр. Гант, Белгия, и е работил като съветник в българската Сметна палата. Той обаче умира рано и издръжката и възпитанието на невръстния Венец Цонев се поема от по-големите му



двама братя юристи, специалисти по данъчно право. Гимназиалното си образование завършва в Немската гимназия в гр. София с отличен успех, а висше образование по право - в Юридическия факултет на Софийския университет също с отличен успех. След изпълнение на задължителния стаж като юрист вместо на правото, той посвещава по-нататъшния си живот на статистическата наука. За това допринася основно преподавателят му по теория на статистиката в Юридическия факултет проф. Димитър Мишайков. Създадените близки контакти и разисквания още от студентските години с проф. Д. Мишайков по проблемите на теорията на статистиката и нейните взаимоотношения с философията и логиката окончателно определят неговата професионална ориентация и съдба. Във връзка с това той приема поканата на ръководителя на катедра „Статистика“ в Софийския университет проф. Д. Мишайков и работи повече от пет години като доброволен асистент по статистика без трудово възнаграждение.

След 1945 г. с обособяването на самостоятелна катедра по статистика към ВИИ „Карл Маркс“¹ (при прехвърлянето на Стопанския факултет от Софийския университет в самостоятелно висше учебно заведение) е назначен за редовен асистент чрез съдействието на новия ръководител на катедрата проф. Иван Стефанов. Както е известно, проф. Ив. Стефанов е утвърден учен статистик, по-късно избран за академик - човек също с много сложна житейска и професионална съдба. Не мога тук да не отбележа, че много пъти в течение на годините по различни поводи проф. В. Цонев е изразявал дълбокото си уважение и респект към двамата професори по статистика - проф. Димитър Мишайков и акад. Иван Стефанов, както по отношение на тяхното научно творчество, така и по отношение на човешките им добродетели и отстояване на академичните принципи. Това се отнася особено до намесата на акад. Ив. Стефанов след зложелателни критични оценки на съдържанието на публикации на проф. В. Цонев, които при тогавашните условия са могли да доведат до сериозни неприятности от идеологически характер. Тези критики се появяват при оценката на различни публикации на проф. В. Цонев през 50-те и 60-те години на миналия век, свързани в една или друга степен с неговите два хабилитационни труда, които не са публикувани - „Върху двукратния корен на приложението на статистиката или статистиката като евристична логика“ (1946 - 1947 г.) и „Върху двукратния корен на статистическата

¹ УНСС.

методология“ (1948 година).² Третият му хабилитационен труд „Диалектика и статистика“, разработен през 1949 г., получава положителен отзив в списанието на Парижката статистическа асоциация.

След повече от 15 години преподавателска работа през 1954 г. проф. В. Цонев е избран за доцент, а през 1964 г. получава званието „професор по статистика“.

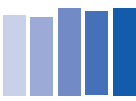
Като преподавател по статистика проф. В. Цонев е подготвил и чел лекции по редица статистически дисциплини. В това отношение могат да се посочат дисциплините „Икономическа статистика“, „Статистика на капиталното строителство“, „Земеделска статистика“, „Вариационна статистика за биолози“ и други. Един от най-значимите статистически курсове, които е подготвил и чел пред студентите от специалност „Статистика“, е курсът по „Планиране и организиране на репрезентативни извадкови изследвания“. Този курс заедно с учебника му „Основи на репрезентативното изучаване“, издаден през 1958 г., се оказва практическо ръководство за много поколения статистици, социолози и други специалисти при проектирането на представителни извадки при най-разнообразни съвкупностни изследвания.

Особено амбициозна задача проф. В. Цонев си бе поставил с разработването и четенето на специален курс по „Логика и стратегия на статистическите изследвания“. В него по нов и изключително оригинален начин се третираха както абстрактно-логически, така и конкретно-практически въпроси на методологията и ефективността на статистическите (съвкупностните) изследвания. Утвърждаването на този курс и лансирането на съдържащите се в него идеи и постановки не беше постигнато в пълна степен поради преждевременното напускане на Института и катедрата по „Статистика“ от проф. В. Цонев през 1977 година³. Тези, които биха проявили интерес към някои от идеите и постановките в този курс, могат да проучат студията му „За необходимостта от разгърнатата теория на съвкупностните изследвания“, поместена в сборника „Изследвания на младежта, теория, методика, практика“, издателство „Народна младеж“, 1978 година.

Тук би трябвало да се посочи и безпрецедентната роля на проф. В. Цонев при разработването и приемането на новия учебен план на специалност „Статистика“ през

² Копия от тези хабилитационни трудове са в досието на проф. В. Цонев в Централния държавен архив, лични фондове, № 1072, заедно с други непубликувани ръкописи.

³ Проф. В. Цонев напуска по морални съображения Института и катедрата „Статистика“, две години преди да има право да се пенсионира. Той е тенденциозно отстранен от Факултетния съвет на Общикономическия факултет поради изразено несъгласие за неправомерно освобождаване от работа заради клеветнически обвинения на преподавател от катедрата „Математика“. Едва през 1990 г. преподавателят беше възстановен на работа в същата катедра.



1975 година. Няма да бъде никак преувеличено, ако се твърди, че той беше в основата на неговата подготовка и главен идеолог при избора, обосновката и подреждането на учебните дисциплини. Този план създаде възможност редица млади преподаватели в катедрата да реализират научния и преподавателския си потенциал. За големия принос на проф. В. Цонев в това отношение може да се съди от факта, че и настоящият учебен план на специалност „Статистика и иконометрия“, разработен при новите условия, съдържа в конструкцията си основни елементи от учебния план от далечната 1975 година.

Заниманията на проф. В. Цонев с проблематиката на планирането и организацията на репрезентативните (представителните) извадкови изследвания го превърнаха в един от най-авторитетните и ерудирани изследователи и приложници в тази все по-развиваща се област през десетилетията у нас. Той ръководи успешно секцията по извадкови изследвания към Научноизследователския институт на Централното статистическо управление през 60-те и 70-те години на миналия век. Под ръководството на проф. В. Цонев и с неговото лично участие се проектираха и изпълниха голяма част от извадковите изследвания в рамките на националната статистическа институция. Тук могат да се посочат:

- ръководство и участие в реорганизацията на извадковите изследвания на домакинските бюджети в страната чрез преход към териториална извадка (един от първите преходи между страните от Източна Европа), както и лансиране на идеята за използване на модела на ротационната извадка;

- ръководство и участие при проектирането на извадковите контролни преброявания на населението за оценяване на систематичната грешка в изчерпателните данни, получени от общите преброявания на населението на страната през 1965 и 1975 година;

- ръководство и участие при проектирането и използването на модела на взаимно проникващите се извадки при изпълнение на т.нар. бързи сводки на материалите от общите преброявания на населението на страната през 1956, 1965 и 1975 година;

- ръководство и участие в разработването на проблемите и моделите на представителни извадки при т.нар. съпътстващи извадкови изследвания при общите преброявания на населението. При преброяването на населението на страната през 1975 г. бяха организирани и изпълнени пет подобни изследвания;

- разработване на т.нар. информационна база за проектиране на обемите на извадките при използване на моделите на гнездовите (кльстърните) извадки при различни типове гнезда за провеждане на демографски и социологически изследвания сред населението;

- ръководство и непосредствено практическо участие при проектирането и внедряването от 1.01.1970 г. на метода на случайните моментни наблюдения за оценка на екстензивното (по време) използване на производствените мощности в над 300 металообработващи и машиностроителни предприятия на страната като текущ, регулярен метод на държавната статистика.

При тези разработки с непосредствено практическо приложение много грижи, внимание и помощ проф. В. Цонев оказва на работата на много млади сътрудници и колеги. Той им импонира със своята страст към съблюдаване на научната истина и уважение към изразяваните мнения и оценки при съвместните дискусии. Това могат да потвърдят много негови сътрудници и колеги. Между тях са чл. кор. проф. Атанас Атанасов, проф. Иванка Съйкова, проф. Петко Божиков, проф. Нигрета Величкова, старши научните сътрудници Мария Динева, Ветка Жекова, Мария Белчева, доц. Андреана Стойкова, авторът на настоящия доклад и много други.

Със завиден авторитет и уважение проф. В. Цонев се ползва сред социологическата общност у нас. Той има сериозен принос за проектирането на емпиричните социологически изследвания (ЕСИ) въз основа на репрезентативните (представителните) извадки, както и при разработването на съществени методологически проблеми, свързани с тези изследвания. Заслугите му за развитието на новосъздадения Институт по социология при проектирането, организацията и изпълнението на конкретни емпирични социологически изследвания са неоспорими. Тук може да се посочи едно от първите ЕСИ, проведено в Института по социология (с директор проф. Живко Ошавков) през 1962 г., за религиозността сред българското население, а също изследването „Градът и селото“. Установеното тясно сътрудничество на проф. В. Цонев с Института по социология намери естествено продължение при участието му в обсъжданията на методологически въпроси от най-различен характер в Центъра за научни изследвания на младежта и главно по проблемите на приложението на репрезентативните извадки, а също и по редица други проблеми на анкетните изследвания. Стенографските материали от обсъжданията на заседанията на т.нар. методологическа група на Центъра дават добра представа за активното участие на проф.



В. Цонев при обсъжданията на различните проблеми от съдържателната работа на Центъра. Редица известни социолози и статистици, работещи в тези среди, са изразявали многократно респекта си към участието на проф. В. Цонев в различните дискусии. Тук могат да се посочат чл.-кор. Ст. Михайлов, проф. Петър-Емил Митев, проф. Йордан Венедиков, доц. Борис Чакалов и много други.

Многостранна и задълбочена е научната продукция на проф. В. Цонев през дългогодишната му творческа работа. Той е автор на близо 80 научни публикации, включващи студии, статии, учебници и рецензии в периодични и други издания. Голяма част са самостоятелни разработки, а останалите - в съавторство. Част от непубликуваните му разработки са на разположение в Централния държавен архив. За отбелязване е, че значителна част от неговите научни разработки (над 15 на брой) са приети и публикувани в чуждестранни статистически и икономически издания в бившите СССР, Чехословакия, Федерална република Германия, ГДР и други.

Прегледът на заглавията и съдържанието на отделните публикации дава представа за насоките и качеството на разработките. Те са резултат на много съвестна и дългогодишна изследователска работа в областта на философията, логиката и съвкупностния подход на изследване на явленията в действителността. Въз основа на богатата си езикова подготовка (свободно владеене на немски, руски, английски и френски език) проф. В. Цонев беше проучил детайлно теоретичните изследвания на почти всички изследователи теоретици от миналото и непрестанно следеше актуалната статистическа литература в световен мащаб. Това му позволяваше да добие енциклопедични познания за развитието на разнообразни идеи в областта на общата теория на статистиката и на теорията на съвкупностните изследвания. В резултат на критичната оценка на вижданията на много наши и чуждестранни автори той съумя да разработи и обоснове редица нови постановки от теоретичен и приложен характер в областта на т.нар. от него „обща теория на статистиката“. Тук не е възможно дори съвсем накратко да се посочат всички теоретични и методологически идеи и постановки, разработени и предложени за дискутиране в неговите публикации. Някои от най-важните сред тях се свеждат до:

- Изясняване на съдържанието на понятието „съвкупност“ като единствен обект при изследване на масовите явления в действителността. Твърди се, че свойствата и тяхната числова оценка на съвкупностите зависят изцяло от т.нар. честотно разпределение на значенията, на признаците на съответните единици.

- Разширяване на съдържанието на понятието „единица“ при прилагане на съвкупностния подход на изследване на масовите явления. Доказва необходимостта да се преодолее представата за физическата природа и физическата обособеност на единиците в изследваните съвкупности. Илюстрира конкретно възможността да се разшири обхватът на тези понятия, като в него се разграничат „парични (конкретно левови) единици“, „елементарни потребителски единици“, съвкупности само с една парична единица и други.

- Разработване на идеята за развитието и преходите между съвкупностите и на особеностите на тяхното изследване чрез съвкупностния подход.

- Поставяне на съвкупностна основа на т.нар. индексна теория, която исторически дълго време се развива откъснато от основния понятиен апарат на общата теория на статистиката. На критична оценка е подложено използването на понятието „разнородна съвкупност“, което е позволило за измерване на динамиката на цените на стоките да се предложат над 120 различни формули конструкции. Обосновани са конкретни решения за конструиране на т.нар. синтетични и аналитични формули при измерване на динамиката и анализа на развитието на съвкупности и на комплекси от съвкупности.

- Разработка на специфичните особености на съдържанието на моделите на обектите при различните видове изследвания: констативно-описателни (ситуационни), обяснителни (диагностични), прогностични, оптимизационни (праксеологични) и на модели с теоретична насоченост. Аргументира се необходимостта от тяхното отчитане с оглед на ефективното планиране и провеждане на съвкупностните изследвания на различните масови явления.

- Разработка на теоретични и методологически постановки за т.нар. от него две нива на изследване на промените (динамиката) в съвкупности: анализ на произточниково и на пофакторно ниво. С числови примери са демонстрирани връзката и различията между двете нива на анализа.

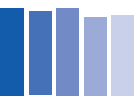
- Забележителен е приносът на проф. В. Цонев при разработката на философските и логическите основи на съвкупностния подход на изследване на масовите явления, за схемите на мисленето в процеса на проектиране, организиране и изпълнение на съвкупностните изследвания. В това отношение той изпъква като уникален учен статистик не само в национален, но и в международен мащаб.

Постиженията и приносите на проф. В. Цонев в теорията и практиката на



съвкупностните изследвания, във вижданията му за обучението на бъдещите специалисти по статистика в УНСС са значими и многопланови. Много от тях не е възможно дори да се посочат в настоящия доклад. Може обаче с основание да се изтъкне, че всяко негово предложение, всяка негова идея за развитието и усъвършенстването на съвкупностния подход на изследване са дълбоко изживени, осмислени и обосновани със съответната аргументация. Очаквам, че когато един ден се преодолеят доста прагматичните сега виждания за теорията на статистиката и на съвкупностния подход при изследване на действителността, творчеството на един бележит български учен статистик ще получи признание, както и непосредствено прилагане в живота. Проф. В. Цонев беше непоправим оптимист, дори романтик по отношение на доброто бъдеще на специалност „Статистика и иконометрия“ и на обучението на пълноценни съвременни специалисти по проектирането и организацията на ефективни съвкупностни изследвания в различни области на действителността.

В заключение искам да изразя моите адмирации към организаторите на настоящата научна конференция от катедра „Статистика и иконометрия“, както и към ръководството на катедрата, че дадоха възможност да се отбележат делата на две забележителни фигури на българската статистическа мисъл през последните 80 години - проф. Стефан Станев и проф. Венец Цонев. Изразявам надеждата, че тяхното творчество няма да бъде забравено от сегашните и бъдещите поколения, работещи в областта на статистиката.



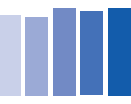
ЦИТИРАНА ЛИТЕРАТУРА:

Алманах „75 години УНСС“. Стопанство, София, 1995, с. 326 - 327.

Мечев, Ст. Проблеми на отрасловите икономически науки в България. ВИИ „К. Маркс“ - София, 1987, с. 528 - 529.

Съйкова, Ив. Честит юбилей, професор Цонев!. Статистика, 2007, кн. 3, с. 12 - 21.

Цонев, В. Моят път като статистик. Статистика, 2009, кн. 1 - 4, с. 228 - 238.



ПРОФ. ВЕНЕЦ ЦОНЕВ - ЕДИН ЗАБЕЛЕЖИТЕЛЕН ТВОРЧЕСКИ ПЪТ В ТЕОРИЯТА И ПРАКТИКАТА НА СЪВКУПНОСТНИТЕ ИЗСЛЕДВАНИЯ

*Тодор Къналиев**

РЕЗЮМЕ В доклада се проследяват жизненият път и творческите постижения на един от най-изявените български учени - статистик и университетски преподавател. Разглеждат се както резултатите от дългогодишната му учебно-преподавателска дейност, така и съществените му постижения в теорията и практиката на съвкупностните изследвания, публикувани у нас и в чужбина. Отбелязва се сериозният му принос в теорията и практиката на представителните извадкови изследвания в страната чрез създадения от него курс и учебник по репрезентативни изследвания. Изтъкнати са приносите му в разработването на връзката между философията, логиката и статистическия (съвкупностния) познавателен процес, използван при различни изследвания.

* Доц. д-р, УНСС, катедра „Статистика и иконометрия“; e-mail: todor_kan@abv.bg.

ПРОФ. ВЕНЕЦ ЦОНЕВ - ВЫДАЮЩИЙСЯ ТВОРЧЕСКИЙ ПУТЬ В ТЕОРИИ И ПРАКТИКЕ ИССЛЕДОВАНИЙ СТАТИСТИЧЕСКИХ СОВОКУПНОСТЕЙ

*Тодор Каналиев**

РЕЗЮМЕ В докладе прослеживается жизненный путь и творческие достижения одного из выдающихся болгарских ученых - статистика и университетского преподавателя. Рассматриваются результаты его многолетней учебно-преподавательской деятельности, а также его значительные достижения в теории и практике исследований совокупностей, опубликованных как у нас, так и за рубежом. Созданные им курс и учебник по репрезентативным исследованиям представляют серьезный вклад в теорию и практику репрезентативных выборочных обследований в стране. Также отмечен его вклад в изучение взаимосвязи между философией, логикой и статистическим познавательным процессом, используемым в различных исследованиях.

* Доц. д-р, УНСС, кафедра „Статистика и эконометрия“; e-mail: todor_kan@abv.bg.



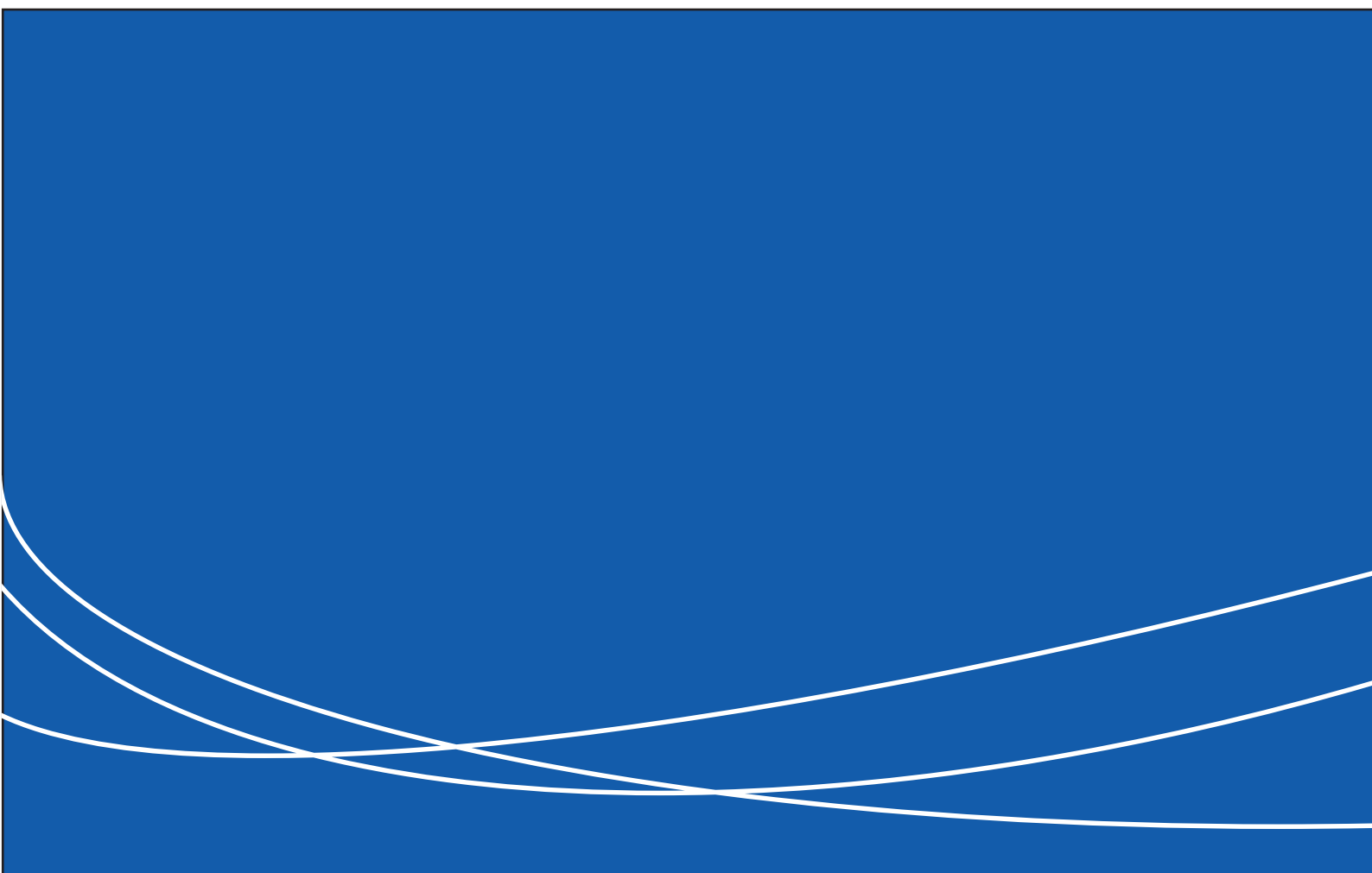
**PROF. VENETS TSONEV - A REMARKABLE CREATIVE PATH IN THE THEORY
AND PRACTICE OF AGGREGATE RESEARCH**

*Todor Kanaliiev **

SUMMARY The paper tracks the life and creative achievements of one of the most prominent Bulgarian scientists - a statistician and a university lecturer. Both the results of his long years of teaching activities, as well as his significant achievements in the theory and practice of the aggregate research published in Bulgaria and abroad, are being considered. Noted is his significant contribution to the theory and practice of representative sample surveys in the country by the created by him course and a textbook on representative researches. His contribution to the development of the relationship between philosophy, logic, and the statistical (aggregate) cognitive process used in various studies has been highlighted.

*Assoc. Prof., UNWE, Department of Statistics and Econometrics; e-mail: todor_kan@abv.bg.

**ИНФОРМАЦИИ, РЕЦЕНЗИИ,
КОНСУЛТАЦИИ**



**„СТАТИСТИКА НА ЦЕНИТЕ НА ЖИЛИЩАТА“ - ПРОМЕНИ В РЕЗУЛТАТ
НА ВНЕДРЯВАНЕ НА ЕДИННА ЕВРОПЕЙСКА РАМКА. ТЕМИ И
НАСОКИ НА РАЗВИТИЕ, ОБСЪЖДАНИ НА МЕЖДУНАРОДНАТА
РАБОТНА СРЕЩА ПРЕЗ 2016 ГОДИНА
(ЛИСАБОН, ПОРТУГАЛИЯ, 28 - 29 АПРИЛ 2016 ГОДИНА)**

*Палмира Фарах**



На 28 и 29 април 2016 г. в Лисабон се проведе ежегодна работна среща на тема „Статистика на цените на жилищата“, на която се извърши оценка на напредъка през изминалата година и се представиха предстоящите задачи и проекти в тази област.

От 2015 г. настъпиха значителни промени в методологията, обхвата и същността на произвежданите в НСИ национални статистически данни за пазарните цени на жилищата. Поради тази причина, преди да продължа с темите, обсъждани на работната среща, ще представя накратко развитието на статистиката в областта на жилищните имоти у нас до момента, промените, настъпили в резултат на нарастващото търсене на сравними данни между държавите в ЕС, и необходимостта от хармонизирана европейска методология за производството им, общи понятия, дефиниции и изисквания по отношение на качеството при измерване на разходите, направени от домакинствата за покупка и стопанисване на жилища.

1. Същност и кратка предистория. Национални и европейски предизвикателства към сектора на жилищните имоти

Статистическото наблюдение за пазарните цени на жилищата започва през 1993 г. и се провежда регулярно до края на 2014 г. на базата на методология, разработена от НСИ.

В периода 1993 - 2014 г. са публикувани следните данни:

- Тримесечни данни
 - Средни пазарни цени на жилища - национално ниво и областни градове
 - Индекси на пазарните цени на жилища, предходно тримесечие = 100, национално ниво, областни градове
- Годишни данни

* Държавен експерт в отдел „Потребителски цени, цени на жилища и паритети на покупателната способност“, дирекция „Макроикономическа статистика“, НСИ; e-mail: pfarah@nsi.bg.

- Средни пазарни цени на жилища - национално ниво и областни градове
- Индекси на пазарните цени на жилища, предходна година = 100, национално ниво, областни градове.

Статистическото наблюдение се е провеждало във фирми, осъществяващи посредническа дейност при сделки с недвижими имоти, където са се регистрирали цени на реални сделки с жилища (апартаменти) между домакинствата - едностайни, двустайни и тристайни апартаменти, с предварително определени параметри.

Наблюдението е провеждано в 27 областни града на страната (NUTS3) с развит пазар на жилища. За периода в тези градове се намират около 54% от жилищата и в тях живее приблизително 50% от населението на страната.

Тъй като наблюдението е обхващало само сделки между домакинствата, извън обхвата остават продажбите на жилища, които са нови за домакинския сектор, т.е. общински или ведомствени жилища, както и новопостроени жилища, чийто брой от 2004 г. насам непрекъснато се увеличава. Само за четири години броят на въведените в експлоатация новопостроени жилища нараства над два и половина пъти - от 8 267 до 20 924 в края на 2008 година.

От друга страна, световната финансова криза през 2007 - 2008 г. засяга в значителна степен жилищния пазар и поставя акцент върху необходимостта от сравнима информация за статистиката на имотите на регионално и европейско ниво.

Преход от национални към хармонизирани данни, произвеждани съгласно единна европейска рамка

Европейската статистическа система, в отговор на нарастващата необходимост от сравними показатели за пазара на недвижимите имоти, изгражда система от индекси на цените на жилища, обитавани от техните собственици (*owner-occupied housing price index - OOH price index*), и индекси на цените на жилищата (*house price indices - HPI*). Разработването и публикуването на тези индекси са от съществено значение за целите на икономическата и паричната политика на държавите - членки на ЕС. Дефлираният индекс за цените на жилищата (дефлиран ИЦЖ) е включен в процедурата за мониторинга на макроикономическите дисбаланси (*Macroeconomic Imbalance Procedure - MIP*) - механизъм за установяване на наличие на макроикономически дисбаланси или прекомерни дисбаланси в държавите членки, при които следва да се предприемат решителни действия за преодоляване на идентифицираните дисбаланси.

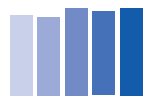
Същевременно и двата индекса са от значение и за домакинствата, тъй като измерват промените в цените на един от най-важните компоненти на разходите на домакинствата и тяхното благосъстояние.

От 1 февруари 2013 г. е в сила Регламент на Европейската комисия № 93/2013 за определяне на подробни правила за внедряването на индексите на цените на жилища, обитавани от техните собственици, в хармонизирания индекс на потребителските цени (ХИПЦ).

Съгласно Регламент 93/2013 България трябва да произвежда два основни набора от данни въз основа на следните разходи, направени от домакинствата:

1. Индекс на цени на жилища, обитавани от техните собственици (ИЦЖОС)

- О.1. Индекс на цени на жилища, обитавани от техните собственици
 - О.1.1. Разходи за придобиване на жилища
 - О.1.1.1. Нови жилища



О.1.1.1.1. Закупуване на нови жилища

О.1.1.1.2. Жилища, построени от самите собственици, и разходи за големи жилищни преустройства

О.1.1.2. Закупуване на съществуващи жилища, които са нови за домакинствата

О.1.1.3. Други услуги, свързани с придобиването на жилища

О.1.2. Разходи на собствениците на жилища

О.1.2.1. Основни ремонти и поддръжка

О.1.2.2. Застраховки, свързани с жилищата

О.1.2.3. Други услуги, свързани със собствеността на жилищата.

2. Индекс на цени на жилища (ИЦЖ)

Н.1. Закупуване на жилища

Н.1.1. Закупуване на нови жилища

Н.1.2. Закупуване на съществуващи жилища.

На 11 май 2016 г. е приет Регламент (ЕС) № 2016/792 на Европейския парламент и на Съвета за хармонизираните индекси на потребителските цени и за индекса на цените на жилищата. С този нормативен акт ИЦЖ и ИЦЖОС са включени в семейството на хармонизираните индекси на потребителските цени заедно с Хармонизираните индекси на потребителските цени (ХИПЦ) и Хармонизираните индекси на потребителските цени при постоянни данъчни ставки (ХИПЦ-ПД). Тази правна рамка е адаптирана към настоящите изисквания и техническия прогрес, като по този начин допълнително подобри приложимостта и съпоставимостта на хармонизираните индекси на потребителските цени и на ИЦЖ.

В отговор на новите предизвикателства и динамичните промени, настъпили на жилищния пазар, през 2008 г. НСИ се включва в поредица от проекти на Европейската комисия и Евростат по разработване на методология и изграждане на система за производство на хармонизирани индекси на цени на жилища. Първото пилотно наблюдение се провежда през четвъртото тримесечие на 2008 година. До края на 2014 г. България изгражда пилотна система за производство на тримесечни индекси на цените на жилища, обитавани от техните собственици.

От 2015 г. в НСИ започва регулярно тримесечно производство на „Индекси на цени на жилища“, с които чрез хармонизиран подход и единен набор от дефиниции се измерва изменението на пазарните цени на жилищата (новопостроени и съществуващи), закупени от български и чуждестранни домакинства на икономическата територия на страната.

ИЦЖ е показател, състоящ се от два подиндекса:

- индекс на цените на нови жилища (новопостроени и вече съществуващи, но нови за сектор „Домакинства“)
- индекс на цените на съществуващи жилища (сделки между домакинствата).

ИЦЖ се изчислява чрез комбинираното прилагане на метода на стратификацията и на метода на хедоник регресията. Наблюдението се провежда в агенции за недвижими имоти, строители и инвеститори в жилищно строителство и общински служби.

От 2016 г. в отговор на непрекъснатото растящите нужди от информация НСИ започва да изчислява и публикува регионални данни за индекси на цени на жилища:

- ИЦЖ по статистически райони
- ИЦЖ за шестте града в България с население над 120 000 жители.

2. Международна работна среща „Статистика на цените на жилищата“

Целта на срещата беше да се анализират постигнатите до момента резултати в конструирането на „Индекси на цени на жилища“ и на „Индекси на обитаваните от собственици жилища“, които държавите членки започнаха регулярно да предоставят на Евростат съгласно Регламент № 93/2013 на Европейската комисия, и да се очертаят основните насоки и приоритети в краткосрочен аспект в зависимост от необходимостта от допълнителни показатели в статистиката на недвижимите имоти. Обхватът на разглежданите теми се разширява извън статистиката на цените на жилища, закупувани от домакинствата, и включва въпроси за производството на допълнителни показатели, отнасящи се за жилищните имоти, като „брой продажби“, „стойност на продажбите“ и др., както и насоки за дефиниране и разработване на индикатори за търговски имоти (*CPPI - Commercial Property Price Indicators*).

В срещата взеха участие представители на държавите - членки на ЕС, Европейската централна банка (ЕЦБ) и Централната банка на Германия (Deutsche Bundesbank). Организатор беше Европейската статистическа служба - Генерална дирекция „Евростат“ на Европейската комисия.

Разгледани бяха следните основни теми:

1. Анализ на резултатите от предоставените съгласно Регламент № 93/2013 данни за ИЦЖ и ИЦЖОС през 2015 г. и представяне на календара за публикуване през 2016 година
2. Инвентаризация (inventory) - инструмент за мониторинг на предоставяните данни (индекси и тегла) за цените на жилищата. Представяне на разработения нов макет на инвентаризация и коментари по предоставената от страните информация за източниците на данни, тегла, методи за обработка и конструиране на индексите, оценки за качество на данните, сравнимост, надеждност и други
3. Методологически въпроси
 - 3.1. Добри практики в осигуряването и обработката на ИЦЖ/ИЦЖОС
 - 3.2. Промени в методологическото ръководство
4. Индекс на цени на търговски имоти (*Commercial property price indices - CPPI*) - практики на някои държави при проучването и изчисляването на *CPPI*
5. Допълнителни индикатори за пазара на жилища (*House sale indicators - HIS*) - наличност на административни данни и определяне на набор от допълнителни индикатори за продажби на жилища.

По точка 1. Анализ на резултатите от предоставените съгласно Регламент № 93/2013 данни за ИЦЖ и ИЦЖОС през 2015 г. и представяне на календара за публикуване на ИЦЖ/ИЦЖОС през 2016 г. г-н Йоаким Ректенвалд, началник на отдел „Статистика на цените; ППС; Статистика на жилищата“ в Евростат, акцентира върху напредъка в областта на производството и публикуването на ИЦЖ през 2015 г. и очерта предстоящите през 2016 г. дейности и срокове за публикуване на индекси и тегла. Подчертано беше, че през 2015 г. значително са подобрени сроковете, в които държавите членки предоставят данни.

През 2015 г. започна регулярното предоставяне в Евростат на индекси на цени на жилища, обитавани от техните собственици. Г-н Ректенвалд направи

оценка на получените данни. Непосредствената цел е разработване на метаданни, както и подобряване на качеството и обхвата на произвежданите ИЦЖ/ИЦЖОС, като се анализират постигнатите резултати и се направи преоценка на методите на компилиране на данните, стратифицирането, източниците на тегла, използването на налични административни данни и други. На 26 май 2016 г. Евростат публикува за първи път индекси на цени на жилища, обитавани от техните собственици.

В момента функционират три канала за разпространение на ИЦЖ/ИЦЖОС и метаданни и е изградена онлайн база данни в сайта на Евростат, в която потребителите имат достъп до индивидуалните тримесечни индекси и метаданни за всяка страна.

Качеството на предоставяните данни и степента на хармонизиране с изискванията на регламента се оценяват на базата на отчети за качеството и метаданни към произвежданите индекси. Държавите членки предоставят и ежегодно актуализират метаданни, които се публикуват на сайта на Евростат.

Мириам Стойра и Робърт Хил от създадената през 2015 г. помощна група (supporting team) в областта на ИЦЖ, ИЦЖОС и на допълнителните показатели за цени на жилища (HSI) направиха подробен преглед на теглата, използвани за изчисляване на индексите. Съществуват големи различия в данните между различните държави в зависимост от използваните източници за тегла (БВП, брой население, жилищен фонд и други). Анализирани бяха причините, които основно са в резултат на разлики в цикличността на жилищния пазар, на социални, исторически и културни различия между страните, както и на различните налични източници на информация и начини на изчисления. Представени бяха сравнителни анализи на теглата, съпоставени с БВП и с БВП на глава от населението, разбити по държави и по категории разходи съгласно Регламент № 93/2013, чл. 1, ал. 1.

По точка 2 „Инвентаризация - инструмент за мониторинг на предоставяните данни (индекси и тегла) за цени на жилища, обитавани от собственици“ г-н Джон Остин от SOGETI запозна присъстващите с резултатите от направения анализ на получените през 2015 г. инвентарни описи за ИЦЖ/ИЦЖОС съгласно новата структура, разработена от SOGETI и представена на предишната работна среща. Последователно и в детайли бяха обсъдени двайсетте основни понятия (concepts) от структурата на инвентаризацията. Дискутираха се проблемни и неясни формулировки, отправени бяха предложения за промени в някои от понятията.

По точка 3 „Методологически въпроси“ представители на шест държави членки направиха преглед на проведените проучвания, изследвания, начини на изчисления и анализ на резултатите за конкретни поиндекси от ИЦЖ/ИЦЖОС.

Г-н Якоб Холмгард от Дания представи разработки на четири метода на регресионен анализ за изчисляване на индекси на цени на жилища: метод на характеристиките, метод с въвеждане на фиктивна променлива „време“ (*time-dummy method*), метод чрез импутации и метод на повторната продажба (*re-pricing method*). Направен беше анализ на предимствата и недостатъците на всеки от методите.

Последваха презентации на Обединеното кралство, Норвегия, Люксембург, Латвия и Полша. Г-н Крис Дженкинс от статистическата служба на Обединеното кралство представи резултатите от нов индекс на цени на жилища, който се изчислява от 2015 г. чрез използването на административни източници на данни. Целта е да се произвежда един официален индекс на цени на жилища, с което да се преодолее основният проблем, свързан с различията в данните между отделните

части на Обединеното кралство (Уелс, Шотландия, Англия и Северна Ирландия) и различните производители на информация - правителство и частни агенции.

Г-жа Мона Тейкл запозна присъстващите с методите на събиране на информация и конструиране на индекса на цените на жилищата в Норвегия. Поради липса на административни източници на данни колегите от статистическата служба на Норвегия провеждат изследване, с което обхващат 50 от най-големите фирми строители на жилищни имоти и агенции за недвижими имоти.

Г-жа Ванда Гуереро от Люксембург представи предизвикателствата, които срещат при конструиране на теглата на ИЦЖ/ИЦЖОС. Основен източник на тегла за ИЦЖОС са данните от национални сметки, докато за тегла на ИЦЖ - данните от Департамента за регистриране на земя и на недвижими имоти. За частта „основен ремонт и поддръжка“ се използват данни за оборота на предприятията от структурната бизнес статистика.

С последвалите презентации на г-жа Анна Бобел и г-жа Звиргждина присъстващите бяха запознати съответно с адаптирането на методологията за производство на ИЦЖОС към условията в Полша и със системата за производство на ИЦЖОС в Латвия.

По втората част от точка 3 г-н Дейвид Фенуик от помощната група в областта на ИЦЖ, ИЦЖОС и допълнителните индикатори за пазара на жилища направи подробен преглед на предложените от тях промени в методологическото ръководство за ИЦЖ/ИЦЖОС. Оживени дискусии предизвика темата за разграничаване на дейностите по „основен ремонт и поддръжка“, които се класифицират към разходите за стопанисване на жилище, от дейностите по „реновация/обновяване“, които се отнасят към разходите за придобиване на жилище. На теория тези понятия са дефинирани ясно с конкретни примери, но практически дейностите трудно могат да бъдат разграничени. Направени бяха предложения в ръководството да се добави таблица с конкретни детайлно описани дейности, отнесени към съответната група разходи.

По **точка 4 „Индекс на цени на търговски имоти“** г-н Ректенвалд представи направеното към момента в тази нова област. В заключителен етап е разработката на методологическо ръководство за индикатори на цени на търговски имоти, източници на данни и методи на изчисления.

Последваха презентации на Португалия и на Испания относно техните практики по отношение на проучване и изчисляване на *CPPI*.

По **точка 5 „Допълнителни индикатори за продажби на жилища“** г-н Ректенвалд акцентира върху обхвата, структурата и техническите процедури по предоставяне на годишни данни за обема на продажбите (стойност на продадените жилища), показател, който 22 държави членки (в т.ч. и България) са започнали да предоставят с набора от данни за годишни тегла за ИЦЖ съгласно т. 5.5 от Регламент № 93/2013. Създадена е работна група „Цени на жилища“, чиято цел е да инициира дискусии между различни международни организации по темата за бъдещото развитие на статистиката на цените на жилищата и свързаните с тях допълнителни индикатори. На първата си среща, инициирана от Организацията за икономическо сътрудничество и развитие през септември 2015 г., работната група е дефинирала три основни групи от показатели, свързани с жилищните имоти:

- Показатели за пазара на жилищните имоти: брой частни домакинства, брой жилища, брой и обем на продажбите на жилища, процент на свободните жилища, начин на ползване (за собствено ползване/за отдаване под наем)
- Показатели за достъпност (свързани с финансовата стабилност): отношение наем/доход, цена/наем, цена/доход



- Други, например индекси на потребителски цени за услуги и продукти, свързани с жилищни имоти, реално нарастване на цените, реално нарастване на кредитите, отношение заем/цена.

С приоритет са индикаторите от първата група и проучването на възможността за предоставяне на детайлни данни по географски признак (населени райони, основни градове, най-големи градове/столица).

Във връзка с това интерес предизвиква презентацията на г-н Грег Патрик от статистическата служба на Ирландия, който запозна участниците със следния набор от допълнителни тримесечни показатели за недвижимите имоти, разработени от помощната група под ръководството на Евростат:

1. Индекс на цени на пазарен наем на жилище
2. Оборот
- 3.а. Индекс на обема на продажбите за всички жилища
- 3.б. Индекс на стойност на продажбите за всички жилища
- 4.а. Индекс на обема на продажбите на нови жилища
- 4.б. Индекс на стойност на продажбите на нови жилища
- 5.а. Индекс на обема на продажбите за стари/съществуващи жилища
- 5.б. Индекс на стойност на продажбите за стари/съществуващи жилища.

От 2012 г. Евростат съфинансира участието на държавите членки за създаването на система за производство или подобряване на съществуващата система за производство на допълнителни индикатори за продажби на жилища с тримесечна периодичност. На първия етап (от 2013 отчетна година) участие вземат Германия и Кипър. През 2014 г. към проекта се присъединяват още девет държави. България с още четири страни се включи на третия етап (от 2015 отчетна година) с дейности по проучване на налични източници за допълнителни индикатори за продажби на жилища.

В заключението си от проведената среща г-н Ректенвал спомена, че пребазирането на ИЦЖ/ИЦЖОС индексите към новата базисна година 2015 = 100 ще стане през юни 2017 г. с предоставяне на данни за първото тримесечие на 2017 година.

Заключения от срещата

- Създадена е единна европейска система от индекси на цени на жилища, обитавани от техните собственици, в съответствие с влезлия в сила през 2013 г. Регламент № 93/2013. Анализът на данните към момента показва някои непълноти и различия между страните. Продължава работата на помощната група (*supporting team*), която предоставя методологическа и практическа помощ на страните с цел преодоляване на различията и хармонизиране на начините и методите за производство на ИЦЖОС.

- През 2016 г. започна регулярно тримесечно публикуване на ИЦЖОС.
- През 2016 г. трябва да се разработят и публикуват метаданни за ИЦЖОС.

- През 2016 г. ще се предложи обновен вариант на инвентаризация (инструмент за мониторинг на предоставяните индекси и тегла на цени на жилища).

- Дефлираният ИЦЖ е един от индикаторите на *Macroeconomic Imbalance Procedure (MIP)*, които се използват за идентифициране на нововъзникващи или устойчиви макроикономически неравновесия в дадена страна. България заедно с

Обединеното кралство и Италия успешно тества разработения макет (*generic template*) на национални отчети за дефлирания ИЦЖ.

- Развитието и динамиката на пазара на недвижими имоти налага необходимост от включване на нови показатели. Разработването на тези допълнителни показатели се осъществява на доброволен принцип от страните и е съфинансирано от Европейската комисия под формата на споразумения за субсидия. От 2015 г. България се включи в третия етап от проекта с дейност „Статистика на недвижимите имоти“, по която отдел „Потребителски цени, цени на жилища и ППС“ ще проучи наличните източници на данни, възможните нива на детайлност и ще разработи методология за производство на пилотна система от допълнителни индикатори за продажби на жилища.

3. Перспективи в областта на статистиката на цените на жилищата в НСИ

В резултат на направените проучвания, анализи и приключилите дейности по споразумение за субсидия от 2015 г. беше създадена пилотна система за тримесечно конструиране на допълнителни показатели за продажби на имоти: „брой продажби“ и „стойност на продажбите“. Източник на данни е имотният регистър, поддържан от Агенцията по вписванията (АВ). В регистъра обаче са включени всички регистрирани в АВ продажби без ясно дефиниране на предназначението на имота: жилищен, търговски, производствен и т.н. Поради това привездането на тази голяма база с данни в използваем формат е едно от предизвикателствата пред отдел „Потребителски цени, цени на жилища и ППС“ в краткосрочен аспект.

През юли 2017 г. Националният статистически институт стартира работата по проект с дейности „Статистика на цените - методологични и практически подобрения и усъвършенствания“ и „Показатели за продажбите на недвижими имоти“ съгласно споразумение за субсидия на Европейската комисия № 04151.2016.005-2016.604. Основната цел по отношение на пазара на жилищните имоти в проекта е свързана с технически и методологически подобрения в областта на показателите за продажби на жилища в България - подобряване на качеството на данните, изчисляване на индикатори за брой продажби на жилища и стойност на продажбите на жилища с тримесечна периодичност, проучване на възможността за използване на административни източници на данни. Предвиждат се срещи с АВ, на които да се търсят начини за подобряване на качеството на имотния регистър, който е единственият изчерпателен източник на данни за продажбите на жилища в страната.

4. Основни понятия и дефиниции

Домакинство означава „домакинство“ по определението в параграф 2.119, букви а) и б) от приложение А към ЕСС 2010, независимо от националността или статута на резидентност.

Икономическа територия на държава членка - територията, определена в параграф 2.05 от приложение А към ЕСС 2010, с изключение на това, че извънтериториалните анклавни, разположени в границите на държавата, се включват в нея, а териториалните анклавни, разположени в останалата част на света, се изключват от нея.

Индекс на цените на жилищата - индекс, който измерва измененията на цените по сделките с жилища, закупувани от домакинствата.



Индекс на цените на жилища, обитавани от техните собственици (ИЦЖОС) - индекс, измерващ изменението на цените по сделките с жилища, които са нови за сектор „Домакинства“, и на цените на други продукти, придобивани от домакинствата в качеството им на обитатели на собствени жилища.

Съпоставимост на хармонизираните индекси - условието индексите на цените на жилища, обитавани от техните собственици, да се считат за съпоставими индекси е всички евентуални разлики между отделните държави на всички равнища на подробност да отразяват само разлики в ценовите промени или в структурата на разходите.

Хармонизирани индекси на потребителските цени - съпоставимите индекси на потребителските цени, които се изготвят от всяка държава - членка на ЕС.

Всички подиндекси на хармонизираните индекси, които се отклоняват от понятията или методите в настоящия регламент, се считат за съпоставими, ако изчислената систематична разлика на получените индекси не превишава един процентен пункт средно за период от една година спрямо предходната година от стойността на индекс, съставен при прилагане на методологическия подход по настоящия регламент - по отношение на ИЦЖОС и ИЦЖ.

**„СТАТИСТИКА НА ЦЕНИТЕ НА ЖИЛИЩАТА“ - ПРОМЕНИ В РЕЗУЛТАТ
НА ВНЕДРЯВАНЕ НА ЕДИННА ЕВРОПЕЙСКА РАМКА. ТЕМИ И
НАСОКИ НА РАЗВИТИЕ, ОБСЪЖДАНИ НА МЕЖДУНАРОДНАТА
РАБОТНА СРЕЩА ПРЕЗ 2016 ГОДИНА
(ЛИСАБОН, ПОРТУГАЛИЯ, 28 - 29 АПРИЛ 2016 ГОДИНА)**

*Палмира Фарах**

РЕЗЮМЕ С информацията се прави кратък преглед на:

- Промените в методологията, обхвата и същността на произвежданите национални статистически данни за пазарните цени на жилищата в НСИ
- Основните теми, обсъждани на работната среща в Лисабон, проведена на 28 - 29 април 2016 година:
 - Анализ на данните за индекси на цени на жилища и за индекси на цени на жилища, обитавани от техните собственици, предоставени съгласно Регламент № 93/2013 на Европейската комисия
 - Инвентаризация (*inventory*) - инструмент за мониторинг на предоставяните данни (индекси и тегла) за цени на жилища
 - Методологически въпроси
 - Индекс на цени на търговски имоти (*Commercial property price indices - CPPI*) - практики на някои държави при проучването и изчисляването на *CPPI*
 - Допълнителни индикатори за пазара на жилища (*House sale indicators - HIS*) - наличност на административни данни и определяне на набор от допълнителни индикатори за продажбите на жилища
- Перспективите в областта на статистиката на цените на жилищата в НСИ.

* Държавен експерт в отдел „Потребителски цени, цени на жилища и ППС“, дирекция „Макроикономическа статистика“, НСИ; e-mail: pfarah@nsi.bg.

**„СТАТИСТИКА ЦЕН НА ЖИЛЬЕ“ - ИЗМЕНЕНИЯ, НАСТУПИВШИЕ В РЕЗУЛЬТАТЕ
ВНЕДРЕНИЯ ЕДИНОЙ ЕВРОПЕЙСКОЙ РАМКИ. ТЕМЫ И НАПРАВЛЕНИЯ РАЗВИТИЯ,
КОТОРЫЕ ОБСУЖДАЛИСЬ НА МЕЖДУНАРОДНОЙ РАБОЧЕЙ ВСТРЕЧЕ В 2016 ГОДУ
(ЛИССАБОН, ПОРТУГАЛИЯ, 28-29 АПРЕЛЯ 2016 ГОДА)**

Палмира Фарах*

РЕЗЮМЕ Краткий обзор по вопросам:

- Изменения в методологии, охвате и сущности национальных статистических данных о рыночных ценах на жилье, исчисляемых в НСИ
- Основные темы обсуждения в рамках рабочей встречи, состоявшейся в Лиссабоне 28-29 апреля 2016 года:
 - Анализ данных об индексах цен на жилье и индексов цен на жилье, занимаемое владельцами, предоставленных Европейской комиссии в соответствии с Регламентом № 93/2013
 - Инвентаризация (*inventory*) - инструмент для мониторинга предоставленных данных (индексы и вес) о ценах на жилье
 - Методологические вопросы
 - Индекс цен на коммерческую недвижимость (*Commercial property price indices - CPPI*) - практика некоторых стран по изучению и расчету CPPI
 - Дополнительные индикаторы рынка жилья (*House sale indicators - HIS*) - наличие административных данных и определение набора дополнительных индикаторов продажи жилья
- Перспективы в области статистики цен на жилье в НСИ.

* Государственный эксперт в отделе „Потребительски цены, цены на жилие и ППС“, дирекция „Макроэкономическая статистика“, НСИ; e-mail: pfarah@nsi.bg.

‘HOUSING PRICE STATISTICS’ - CHANGES AS A RESULT OF THE IMPLEMENTATION OF A SINGLE EUROPEAN FRAMEWORK. THEMES AND GUIDELINES FOR DEVELOPMENT DISCUSSED ON THE INTERNATIONAL WORKING MEETING IN 2016 (LISBON, PORTUGAL, 28-29 APRIL 2016)

*Palmira Farah**

SUMMARY This information is briefly reviewing:

- Changes in the methodology, scope and nature of the produced national statistics on housing market prices in the NSI
- Main topics discussed at the workshop in Lisbon on 28-29 April 2016:
 - Analysis of housing price indices and price indices of owner-occupied dwellings provided under European Commission Regulation 93/2013
 - Inventory - a tool for monitoring the data provided (indices and weights) for housing prices
 - Methodological issues
 - Commercial property price indices (CPPI) - practices of some countries in the study and calculation of CPPI
 - Additional Market House Indicators (*House sale indicators - HIS*)
 - availability of administrative data and determination of a set of additional indicators for housing sales
- Prospects in the field of housing price statistics in the NSI.

* State Expert in Consumer Prices, Housing prices and PPP Department, Macroeconomic Statistics directorate, NSI; e-mail: pfarah@nsi.bg.

НА ВНИМАНИЕТО НА АВТОРИТЕ

Статиите, предложени на редакцията на сп. „Статистика”, трябва да бъдат оригинални, непубликувани в други издания.

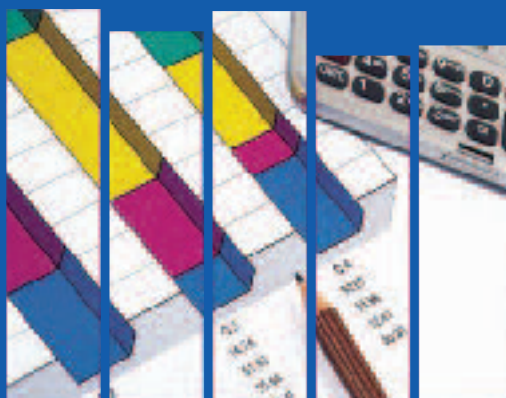
По решение на Редколегията на авторите се предоставя възможност статиите им да бъдат публикувани или на български, или на английски език. За целта е необходимо авторът, който желае статията му да бъде публикувана на английски език, да представи в редакцията ръкописи и съответни файлове и на български, и на английски език при спазване на посочените изисквания и за двата ръкописа. Авторите трябва да имат предвид, че английският текст ще бъде публикуван без редакционна намеса, което означава, че те носят пълна отговорност за коректността на превода.

Статиите, придружени от кратко резюме на български и на английски език и номерата на научната област от JEL класификацията, се представят в редакцията на сп. „Статистика”. Текстът трябва да бъде на Word, а графиките - Excel, със задължително приложени данни. Формулите в текста и тези на самостоятелен ред (без числата в тях) трябва да бъдат задължително курсивни (Italic). Ако графиките са дадени отделно, в текста трябва да се посочи точното им място. Авторите носят отговорност за коректността на данните и цитатите. Ръкописите трябва да са придружени от трите имена, научната степен и звание, точния адрес, телефона, длъжността, мястото на работа и електронния адрес (e-mail) на автора.

Поместването на статия на страниците на списанието не означава, че предложенията и евентуалните критични бележки, които авторът отправя, се споделят от Редакционната колегия или от ръководството на Националния статистически институт.

Редакцията на сп. „Статистика” би приветствала различни инициативи на читателската аудитория под формата на дискусии на страниците на списанието, научно-професионални „реплики” по повод на отпечатана статия, самостоятелни кореферати по повдигнатата тема - изобщо всяка инициативност по теми и проблеми на науката и практиката на статистическите изследвания от сферата на българската и европейската статистика.

Редакция на списание „Статистика”



**СТАТИСТИКА
STATISTICS**

1/2018

www.nsi.bg