

## ВЛИЯНИЕТО НА ИНДЕКСА НА ТЕЛЕСНАТА МАСА И ТЮТЮНОПУШЕНЕТО ВЪРХУ РАЗЛИКАТА В СРЕДНАТА ПРОДЪЛЖИТЕЛНОСТ НА ПРЕДСТОЯЩИЯ ЖИВОТ

*Петя Брайнова\**

### 1. Въведение в основните декомпозиционни методи

#### 1.1. Въведение

Настоящото изследване има за цел да представи един известен детерминистичен метод, наречен **Data Envelopment Analysis (DEA)**, и да покаже неговото приложение в областта на демографската статистика. Класическото предназначение на DEA е да сравнява хомогенни единици на базата на тяхната ефективност. Той пресмята ефективността на всяка единица като отношение между произведената продукция към изразходваните за нейното производство ресурси. В настоящия анализ **сравнителните единици** са 12 възрастово-полови групи (6 за мъжете и 6 за жените: 18 - 24, 25 - 34, 35 - 44, 45 - 54, 55 - 64, 65 - 74). **Производенията** е сумата от вече изживените среден брой години на всяка група и среден брой години предстоящ живот. **Ресурсите** са свързани със здравословния начин на живот на всяка група, а именно относителните дялове на непушачите и на хората с **индекс на телесната маса (ИТМ)** в норма<sup>1</sup> (приложение, табл. 1). Хипотезата е, че при увеличаване на дяловете на лицата с ИТМ в норма и на непушачите, средната продължителност на предстоящия живот също нараства в резултат на здравословния начин на живот, който води до дълголетие. По този начин DEA влиза в ролята на декомпозиционна техника на разликата в средната продължителност на живота между мъжете и жените и резултатът от това е, че за всяка от шестте възрастови групи ще може да се каже колко голяма би била изследваната разлика, ако продължителността на живота зависеше само от въздържането от тютюнопушене или само от наличието на ИТМ в норма.

Основно предимство на използвания подход е, че той няма ограничения за броя на ресурсите (както и за броя на продуктите) и при наличие на данни може да се приложи за повече от два фактора, характеризиращи здравословния начин на живот. Настоящото изследване има за цел по-скоро да онагледи приложението на DEA в демографската област, а не толкова да бъде изчерпателно по отношение на факторите. Затова се ограничавам само до два от тях с

\*Докторант, катедра „Статистика и иконометрия”, УНСС - София; e-mail: petia\_brainova@abv.bg.

<sup>1</sup> Норми на ИТМ според Световната здравна организация, 1995 г. (World Health Organization (WHO): ИТМ под 18.5 - поднормено тегло, ИТМ от 18.5 до 24.9 - нормално тегло, ИТМ от 25 до 29.9 - наднормено тегло, ИТМ от 30 до 34.9 - затлъстяване (клас 1), ИТМ от 35 до 39.9 - затлъстяване (клас 2), и ИТМ 40 или повече - болестно затлъстяване.

цел опростяване на декомпозиционната формула<sup>2</sup>, по-кратка интерпретация и по-прегледно изобразяване на резултатите. Настоящото приложение на DEA в демографската статистика е ново, нетрадиционно и идеята за него е изцяло моя.

## 1.2. Декомпозиционни техники

Декомпозиционните техники са едни от основните постижения в демографската статистика от втората половина на XX в. до днес. Kitagawa (1955) извежда по математически път формула за декомпозиране на разликата между два дяла според влиянието на един фактор. При два фактора, освен ефектите на всеки от факторите, формулата на Kitagawa (1955) съдържа ефект на взаимодействието между тях, който е труден за интерпретация. Двадесет години по-късно Retherford и Cho (1973) оптимизират формулата, като успяват да отстранят ефекта на взаимодействието. Проблемът е, че техният метод дава различни резултати в зависимост от реда на пресмятане на ефектите от различните фактори. Няколко години по-късно друг талантлив учен - Das Gupta (1978), успява да изведе декомпозиционна формула без ефект на взаимодействие и същевременно без ограничение от реда на пресмятане на отделните ефекти. През следващите няколко години Das Gupta (1989, 1991, 1993) продължава да работи по същата тема.

## 1.3. Декомпозиция на средната продължителност на предстоящия живот

През 80-те години на XX в. започват първите опити за приложение на декомпозиционните техники върху средната продължителност на живота  $e_x$ . На базата на техниката на Kitagawa (1955) United Nations (1982) предлага декомпозиция на три възрастови групи, а именно 0 - 29, 30 - 64 и 65+. В същия период са създадени редица методи за декомпозиране на разликите в средната продължителност според допринасянето на различните нива на смъртност по възраст (Arriaga (1982, 1984), Pollard (1982, 1988), Pressat (1985)). Методите на Arriaga (1984) и Pollard (1982, 1988) работят на един и същи принцип, като основната разлика между тях е, че Arriaga (1984) използва директно дискретните данни от таблиците за смъртност, а Pollard (1982, 1988) работи с техни непрекъснати приближения, базирани на интензивността на смъртността в безкрайно малък възрастов интервал. Идеята за непрекъснатост е заложена още в метода на Keyfitz (1977), който формулира зависимостта между две логаритмични производни на времето - едната на интензивността, а другата на средната продължителност на предстоящия живот при раждане. През същата година Русев (1977) извежда формула за зависимостта на средната продължителност на живота от повъзрастовата смъртност. Vaupel (1986) - един от

<sup>2</sup> Броят на събираемите в декомпозиционната формула нараства правопрпорционално на броя на използваните фактори (ресурси).



последователите на Keyfitz, предлага непрекъсната декомпозиция по средния брой изгубени години, приемайки ограничението, че смъртността се променя пропорционално за всички възрасти. Castro (2001) прави сравнителен анализ на методите на United Nations (1982), Arriaga (1984), Pollard (1982, 1988) и Vaupel (1986), като изследва техните предимства и недостатъци, както и случаите, в които всеки от тях работи най-добре. Редица български автори също имат принос в анализа на възрастовата смъртност и средната продължителност на живота. Методът на Arriaga е разгледан и приложен за първи път у нас от Жекова (2009). Две години по-късно Жекова (2011) прилага анализа на съответствието (correspondence analysis), за да изследва **пространствената асиметрия във възрастовата смъртност**<sup>3</sup> на населението в България. Анализът позволява да се отграничат както териториалните единици, които детерминират пространствената асиметрия, така и възрастовите групи, които допринасят в най-голяма степен за нейното формиране. Калоянов (2011) прилага метода на Arriaga (1984) върху данните за България и посредством него изследва динамиката на средните продължителности на живот на мъжете и жените поотделно за периода 1974 - 2009 г., както и тази на различията между тях за двата периода 1974 - 1976 и 2007 - 2009 година.

#### **1.4. Декомпозиция на средната продължителност на предстоящия живот по причини за умиране**

През 80-те години анализът на средната продължителност на живота продължава в друга посока - декомпозицията е пренесена върху различните причини за умиране. Методите на Arriaga (1982, 1984), Pollard (1982), Pressat (1985) са приложими и при изследване на причините за умиране, а Pollard (1988) показва, че тези методи са еквивалентни помежду си и също така еквивалентен на тях е дискретният метод на Andreev (1982). Preston и Beltran-Sanchez (2007) извеждат алтернативна декомпозиционна формула за анализ по причини на Arriaga (1984), използвайки метода на Kitagawa (1955). Основната разлика между подходите на Preston и Beltran-Sanchez и Arriaga (1984) е, че при Preston и Beltran-Sanchez (2007) приносът на всяка отделна причина е претеглен със средната вероятност за доживяване при условие, че действат всички причини, с изключение на тази.

#### **1.5. Обобщени алгоритми за декомпозиция**

Andreev, Shkolnikov и Begun (2002) предлагат алгоритъм за декомпозиция на разликата в стойностите на обобщени демографски показатели по възраст и други фактори. Този алгоритъм представлява постъпков метод за преобра-

<sup>3</sup> Съществени отклонения при част от наблюдаваните териториални единици по отношение на нормално съществуващата териториална вариация около средните равнища на възрастова смъртност.

зуване (stepwise replacement) на матрица с елементи, обобщени демографски показатели в матрица, съдържаща същия тип елементи, но отнесени за друга съвкупност (или същата съвкупност, но към друг времеви момент). Примерни демографски показатели са средната продължителност на живота както в класическия си вариант, така и в добро здраве, т.нар. health expectancy. Този алгоритъм за постъпково заместване се явява обобщена версия на съществуващите методи за декомпозиция на разликата в средната продължителност на живота. Съществени постижения в анализа на непрекъснатата променлива имат Vaupel и Romo (2002, 2003), които извеждат, че първата производна на  $e_0(t)$  по времето  $t$  е интеграл от произведението на степента на понижаване на смъртността  $\rho(x,t)$ , оставащата продължителност на живота  $e(x,t)$  и вероятността за умирање  $f(x,t)$  по възраст. Също през 2003 г. Romo (2003) публикува обзорната си разработка, посветена на декомпозиционните методи в демографията, където описва най-значимите от тях и ги съпоставя един с друг.

## 2. Европейско здравно интервю. Описание на използваните данни

### 2.1. Методология на извадката на Европейското здравно интервю

През периода 2008 - 2010 г. според официалните данни на Националния статистически институт (НСИ) разликата в средната продължителност на предстоящия живот между мъжете и жените на възраст 18 години в България е малко повече от седем години<sup>4</sup>. През 2008 г. Евростат е провел представително социологическо изследване на домакинствата, наречено Европейско здравно интервю (European Health Interview Survey (EHIS), по метода интервю **лице в лице**. Планирано е изследването да се провежда на всеки пет години. Събраната количествена информация по пол, възрастови групи, страни (в т.ч. България) и други характеристики е със свободен достъп, предоставен от Евростат.

За България изследването е проведено от НСИ, като в съответствие с методологията на Евростат, то е представително за домакинствата в страната, с изключение на институционализираните домакинства<sup>5</sup>. Извадката е двустепенна гнездова, районирана по административни области и местоживеене (град/село). В резултат на това са формирани 56 страти. На първа степен са избрани 745 гнезда (населени места), а на втора степен във всяко от тях чрез систематичен подбор са избрани по пет домакинства - общо 3 725 домакинства. За интервю са поканени всички лица на 15 и повече години в тези домакинства. Събрана е информация за 5 661 лица, които са 73.8% от предварително избраните 8 393 души, като ненамерените или отказалите интервю не

<sup>4</sup> Официалните данни са дадени за всички последователни възрасти от 0 до 100 години, като изчисленията на средните продължителности на различните възрастови групи са осъществени от автора.

<sup>5</sup> В т.ч. пансионни и общежития, институции за предоставяне на социални услуги, затвори и т.н.



са били подменяни. Заложената максимална абсолютна грешка за 8 393 лица при относителен дял 50% е 1.27%.

## 2.2. Избор на здравни фактори

Обобщените резултати от изследването представят данни за здравния статус на населението, т.нар. „детерминанти на здравето“ (Determinants of health) и други фактори, свързани със здравето, като:

- Дълготрайни заболявания или други здравни проблеми
- Субективна оценка на здравния статус
- Дейности, прекратени поради здравословни проблеми
- Тютюнопушене
- Употреба на алкохол
- Ръст и тегло, на базата на които се изчислява ИТМ<sup>6</sup>.

В настоящото изследване са използвани относителните дялове на лицата с ИТМ в норма и на непушачите. Причините за избора на точно тези два фактора, характеризиращи здравословния начин на живот, са няколко:

- Те са две от шестте „детерминанти на здравето“ според Европейското здравно интервю.

- ИТМ е стандартен медико-биологичен показател, който служи за определяне на здравословното тегло при определен ръст и за диагностициране на затлъстяване и недохранване (Център за контрол на болести и превенция (CDC).

- При мъжете еднофакторният линеен модел представя основните компоненти на връзката между независимата променлива „относителен дял на непушачи“ и зависимата променлива „общ брой години живот“. Същият извод важи и за жените, като стойността на емпиричната характеристика  $F_{1,eu}^7$  при тях е по-малка от тази при мъжете, но все пак достатъчно голяма, за да се отхвърли нулевата хипотеза  $H_0$  при  $\alpha = 0.01$ .

- Разликата в относителните дялове на лицата с ИТМ в норма между мъжете и жените по възрастови групи (по данни от Европейското здравно интервю) приема както положителни, така и отрицателни стойности (приложение, табл. 1), което благоприятства тестването на разнопосочни ефекти при декомпозицията.

<sup>6</sup> Индексът на телесната маса или Body mass index (BMI) се определя от съотношение на теглото към квадрата на ръста и съгласно метричната система (СИ) се измерва в  $\text{kg}/\text{m}^2$ .

<sup>7</sup> Поради специфичния групиран характер на данните нормалните уравнения за намиране на параметрите на регресионното уравнение са предварително модифицирани. Пресметнати са отделни регресионни модели за мъжете и жените. Отчетени са теглата на възрастовите групи, чиито стойности след модифицирането са именно относителните дялове непушачи в съответната група.

### 2.3. Ограничения на данните

Избраните възрастови групи не обхващат цялото население и първата от тях обхваща по-кратък период (7 години) от останалите (10 години). Причините за това произхождат от естеството на социологическите данни и самите здравни фактори. Наличните данни за тютюнопушенето се отнасят за населението на 15 и повече години<sup>8</sup>, а тези за ИТМ - за пълнолетното население<sup>9</sup>. При интерпретацията на ИТМ трябва да се има предвид, че този индекс **не работи** добре при някои специфични подгрупи от населението, а именно:

- бременните жени;
- активно спортуващите, атлетите, спортистите и т.н.;
- най-възрастните хора (горната граница на възрастта може да варира по ред причини).

Причината за ограничената употреба на ИТМ произхожда от неговата дефиниция. Тя взема предвид само стойностите на теглото и ръста на отделния индивид, но не и съотношението между мускулна и друга маса. Освен това, при някои от горепосочените групи по-висок ИТМ невинаги е предпоставка за по-лошо здравословно състояние (напр. бременните) (Център за контрол на болести и превенция (CDC)).

## 3. Използван метод

### 3.1. Описание на използвания метод

Декомпозицията на разликата в средната продължителност на предстоящия живот е осъществена по метода DEA (Data Envelopment Analysis), който е базиран на техниката на линейното оптимиране. В една стандартна линейна оптимизационна задача целта е да се максимизира/минимизира, т.нар. **целева функция**, при определени ограничения. Целевата функция представя линейна комбинация на изследваните променливи, като неизвестните в задачата са техните коефициенти, играещи ролята на своеобразни **тегла**. Върху неизвестните е наложена система от ограничения и задачата е да се намерят стойностите на теглата, които оптимизират целевата функция при тези ограничения. Решението се получава по известния **симплекс алгоритъм** - итеративен метод, който обхожда всички възможни решения, намиращи се на границата на областта от допустимите решения, докато намери оптималното решение.

Декомпозиционният анализ е представен във вид на линейна оптимизационна задача. Изчисленията по оптимизационната задача са осъществени посредством екселско приложение на DEA, което по зададени входни данни

<sup>8</sup> За целите на изследването правя допускането, че непушачите са разпределени равномерно по възраст в групата 15 - 24 години, за която се отнасят данните на Евростат, и затова можем да ползваме същия дял за групата 18 - 24 години.

<sup>9</sup> ИТМ може да се изчисли и за лица под 18 години, но това става по модифицирана формула.



автоматично, на базата на **симплекс метода**, пресмята оптималните тегла, наречени **множители**. DEA е метод, който сравнява еднородни единици на базата на измерване на тяхната **ефективност** като съотношение на продукцията и използваните за нейното производство ресурси. Всяка от възрастово-половите групи е в ролята на **еднородна сравнителна единица**, като общо разгледаните единици са 12 - по 6 възрастови групи за всеки пол. Сумата на изживените до момента години и предстоящите години е разгледана в качеството на **продукт** от здравословния начин на живот, характеризирани от **ресурсите** наличие на ИТМ в норма и въздържане от тютюнопушене. DEA предлага множество модели в зависимост от естеството на задачата, която трябва да се реши. Те се характеризират по своята **ориентация** и **брой на оптималните нива** на ефективност. Cooper, Seiford и Tone (2007) описват подробно различните DEA модели в едноименната си книга, както и множество примери за приложението и спецификата на всеки от тях. Cook и Zhu (2008) публикуват практическо ръководство с геометричните интерпретации на основните DEA модели, което до голяма степен улеснява разбирането на методологията.

### 3.2. Характеристики на използвания модел

Използваният DEA модел се нарича **Input-oriented constant return to scale (CRS) model with restricted multipliers**, т.е. ориентиран към входа модел с ограничени множители, който допуска едно ниво на ефективност за оптимално. В ориентирания към входа модел ефективността на всяка изследвана единица приема стойности от 0 до 1 (0 - 100%). Тази стойност показва какъв процент от използваните ресурси биха оправдавали получената продукция, така че производственият процес да бъде ефективен. При модела, ориентиран към входа, нивата на произведения продукт се считат за постоянни, а нивата на използваните ресурси влизат в ролята на променливи. Стойността на ефективността е отговорът на въпроса: **Ако се запази същото ниво на продукт на изхода, каква част от ресурсите на входа е резонно да се използват, за да бъде достигната границата на производствените възможности?**

Освен с ориентацията си, в случая към входа, моделът се характеризира и с броя на оптималните си нива на ефективност. Използваният модел определя едно-единствено ниво като 100% ефективно и съпоставя всички единици с него. Възможно е няколко единици да получат 100% само ако имат ефективности, равни както помежду си, така и равни на оптималната разглеждана ефективност. Вследствие неефективните трябва да догонят съотношението на продукция-ресурси на първенеца/първенците, за да станат ефективни като него/тях, а ефективните са еталон за подражание. CRS моделът е подходящ за използване, когато всички включени в модела единици са еднородни не само по вид на произвежданата продукция, но и по мащаб на производството.

### 3.3. Аналитичен вид на ефективността

За всяка сравнителна единица ефективността представлява съотношение на продукцията и използваните за нейното производство ресурси или по-точно претеглената сума на продуктите към претеглената сума на ресурсите, където теглата, наречени **множители**, са неизвестните в модела:

$$Efficiency = \frac{\sum w_i * product_i}{\sum w_j * resource_j}. \quad (1)$$

В конкретния случай ефективността за всяка от 12-те възрасто-полови групи има вида:

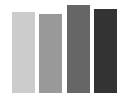
$$Efficiency_i = \frac{w_i^{years} * total\_life\_years_i}{w_i^{BMI} * \%BMI_i + w_i^{no-sm} * \%No\_smoke_i}, \quad (2)$$

където  $i$  се мени от 1 до 12 и значението на променливите е:

- $\%BMI_i$  - дялът на хората с ИТМ в норма в  $i$ -тата възрасто-полова група;
- $w_i^{BMI}$  - неизвестното тегло (множител) на ресурса „ИТМ в норма” в  $i$ -тата възрасто-полова група;
- $\%No\_smoke_i$  - дялът на непушачите в  $i$ -тата възрасто-полова група;
- $w_i^{no-sm}$  - неизвестното тегло (множител) на ресурса „въздържане от тютюнопушене” в  $i$ -тата възрасто-полова група;
- $total\_life\_years_i$  - сумата на вече изживените години и средния брой предстоящи години живот на  $i$ -тата възрасто-полова група;
- $w_i^{years}$  - неизвестното тегло (множител) на продукта „общ брой години живот” в  $i$ -тата възрасто-полова група.

При модела, ориентиран към входа, претеглената сума на ресурсите се намира в знаменател, а тази на продуктите - в числител. Възрасто-половите групи, които са в ролята на сравнителни единици, са еднородни не само по вид на произвежданата продукция, но и по мащаб на производство, тъй като за продукт е взет общият брой години живот, а не само средната продължителност на предстоящия живот на всяка група. Последното обосновава използването на модел с едно оптимално ниво на ефективност. За всяка от 12-те групи ДЕА търси стойностите на неизвестните тегла, така че ефективността да бъде най-голяма, което при фиксирана стойност на продукта (общия брой години живот) е еквивалентно на търсене на най-малката стойност на претеглената сума на ресурсите (% лица с ИТМ в норма и % непушачи).





### 3.4. Ограничаване на допустимите решения

Това, че избраният модел е с **ограничени множители**, означава, че оптимизирането се прави при зададени интервали, в които могат да варират съотношенията на продуктите и ресурсите. Последното е допълнително изискване, което е добре да се използва при наличие на необходимата информация. В нашия случай поради представителността на данните приемаме, че разглежданите 12 единици изчерпват всички възможни единици, които осъществяват процеса на **производство** на жизненни години, **влагайки** в него здравословен начин на живот, измерен с ИТМ в норма и въздържане от тютюнопушене през 2008 година. В случая има налични данни за пресмятането на съотношенията между продукта и всеки от ресурсите, както и между двата ресурса за всяка възрастово-полова група. Получените съотношения варират в определени числови интервали, които могат да се използват за дефиниране на ограниченията на модела и по-този начин да се стесни множеството на допустими решения на задачата.

## 4. Връзка между използвания метод и декомпозицията

### 4.1. Интерпретация на ДЕА множителите

В резултат на прилагането на ориентирания към входа модел за всяка една от 12-те възрастово-полови групи се получават четири числа: стойността на ефективността (число между 0 и 1) и оптималните стойности на търсените тегла (множители) на продукта и на двата ресурса (приложение, табл. 2). Оптималният множител на всеки ресурс показва влиянието на неговата значимост върху обема на продукцията в сравнение с това на останалите ресурси. Поради това в рамките на една възрастово-полова група оптималните множители на ресурсите се явяват сравнителни характеристики на степените им на влияние върху продукцията. Например ако за  $i$ -тата група е вярно, че  $w_i^{no-sm} \approx 2 * w_i^{BMI}$ , това означава, че в тази група влиянието на ИТМ върху общия брой години живот е приблизително два пъти по-слабо от това на нетютюнопушенето. С други думи, ДЕА остойносттава степента на влияние на двата фактора за всяка от групите, с което може да се декомпозира по тях.

### 4.2. Декомпозиция на общия брой години живот ( $total\_life\_years_i$ )

При изчисляването на ефективността ориентираният към входа модел приравнява на единица претеглената сума на ресурсите и от формула (2) следва, че ефективността е равна на претегления продукт. Откъдето следва, че:

$$total\_life\_years_i = \frac{Efficiency_i}{w_i^{years}} \quad (3)$$

Като използваме, че теглата на ресурсите удовлетворяват равенството:

$$w_i^{BMI} * \%BMI_i + w_i^{no-sm} * \%No\_smoke_i = 1, \quad (4)$$

получаваме следната зависимост:

$$total\_life\_years_i = \frac{Efficiency_i}{w_i^{years}} * w_i^{BMI} * \%BMI_i + \frac{Efficiency_i}{w_i^{years}} * w_i^{no-sm} * \%No\_smoke_i. \quad (5)$$

Последното равенство може да се интерпретира като декомпозиция на общия брой години живот в  $i$ -тата възрастово-полова група. Например при  $i=1$ , тази зависимост се отнася за мъжете във възрастовата група (18 - 24 години). Първото събираемо показва колко от общия брой години живот се дължат на наличието на ИТМ в норма, а второто събираемо - на въздържането от тютюнопушене. Тази интерпретация е чисто абстрактна и е коректна единствено при условие, че сме изолирали влиянието на всички останали фактори, които определят продължителността на живота. Именно поради тази причина преминаването на декомпозицията от ниво „възрастово-полова група“ към ниво „възрастова група“ е от съществено значение за практическата стойност на интерпретацията на резултатите. На ниво „възрастова група“ обектът на декомпозиция е самата разлика в средната продължителност на предстоящия живот между мъжете и жените в рамките на тази възрастова група, което е и основният фокус на разглеждания анализ.

#### 4.3. Декомпозиция на разликата в средната продължителност на живота ( $life\_exp\_diff_{(AGE\ GROUP)}$ )

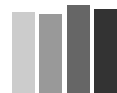
По отношение на една и съща възрастова група разликата в средната продължителност на живота между мъжете и жените -  $life\_exp\_diff_{(AGE\ GROUP)}$ , се получава като разлика между общия брой години живот на мъжете и жените в съответната група. Например за най-младата група - 18 - 24 години:

$$life\_exp\_diff_{(18-24)} = total\_life\_years_7 - total\_life\_years_1, \quad (6)$$

т.е. от общия брой години живот на жените на възраст 18 - 24 години изваждаме този на мъжете и тъй като изживените вече години и при двете съвкупности са по 18 години<sup>10</sup> се получава разликата в средните им продължителности на предстоящия живот на възраст 18 години. Замествайки формула (5) в равенство (6), получаваме:

$$life\_exp\_diff_{(18-24)} = \left[ \frac{Efficiency_7}{w_7^{years}} * w_7^{BMI} * \%BMI_7 - \frac{Efficiency_1}{w_1^{years}} * w_1^{BMI} * \%BMI_1 \right] + \left[ \frac{Efficiency_7}{w_7^{years}} * w_7^{no-sm} * \%No\_smoke_7 - \frac{Efficiency_1}{w_1^{years}} * w_1^{no-sm} * \%No\_smoke_1 \right]. \quad (7)$$

<sup>10</sup> При отчитането на изживените години е използвана долната граница на възрастовия интервал и съответно средният брой предстоящи години също са изчислени от долната граница нататък. Друг възможен подход е взимането на средите на възрастовите интервали.



Всяко от събираемите във формула (7) се състои от три множителя. За да се отделят онези от тях, които съответстват на декомпозиционните фактори, се прилага известната формула на Kitagawa (1955) за всеки от изразите в квадратните скоби и се използват зависимостите (3) и (6):

$$\begin{aligned}
 & life\_exp\_diff_{(18-24)} = \\
 & \left[ \begin{aligned}
 & life\_exp\_diff_{(18-24)} * 0,5 * (w_7^{BMI} * \%BMI_7 + w_1^{BMI} * \%BMI_1) + \\
 & + 0,5 * (total\_life\_years_7 + total\_life\_years_1) * \\
 & * (w_7^{BMI} * \%BMI_7 - w_1^{BMI} * \%BMI_1)
 \end{aligned} \right] + \\
 & \left[ \begin{aligned}
 & life\_exp\_diff_{(18-24)} * 0,5 * (w_7^{no-sm} * \%No\_smoke_7 + w_1^{no-sm} * \%No\_smoke_1) + \\
 & + 0,5 * (total\_life\_years_7 + total\_life\_years_1) * \\
 & * (w_7^{no-sm} * \%No\_smoke_7 - w_1^{no-sm} * \%No\_smoke_1)
 \end{aligned} \right]. \quad (8)
 \end{aligned}$$

Формула (8) представлява декомпозиция на разликата в средната продължителност на живота между мъжете и жените по два фактора - в случая наличие на ИТМ в норма и въздържане от тютюнопушене. Изразът в първите квадратни скоби показва ефекта от наличието на ИТМ в норма, а изразът във вторите квадратни скоби - ефекта от въздържане от тютюнопушене. За улеснение правим полагането:

$$\begin{aligned}
 A & \doteq life\_exp\_diff_{(18-24)} * 0,5 * (w_7^{BMI} * \%BMI_7 + w_1^{BMI} * \%BMI_1), \\
 B & \doteq 0,5 * (total\_life\_years_7 + total\_life\_years_1) * (w_7^{BMI} * \%BMI_7 - w_1^{BMI} * \%BMI_1), \\
 C & \doteq life\_exp\_diff_{(18-24)} * 0,5 * (w_7^{no-sm} * \%No\_smoke_7 + w_1^{no-sm} * \%No\_smoke_1), \\
 D & \doteq 0,5 * (total\_life\_years_7 + total\_life\_years_1) * \\
 & * (w_7^{no-sm} * \%No\_smoke_7 - w_1^{no-sm} * \%No\_smoke_1).
 \end{aligned}$$

След полагането формула (8) придобива вида:

$$life\_exp\_diff_{(18-24)} = [A + B] + [C + D]. \quad (9)$$

#### 4.4. Интерпретация на различните ефекти

Събираемото  $A$  от формула (9) показва колко голяма би била разликата  $life\_exp\_diff_{(18-24)}$ , ако върху нея влияе само факторът ИТМ и ако това влияние се осъществява в еднаква степен при мъжете и жените на възраст 18 - 24 години, а именно тяхната средна аритметична стойност. Ако искаме да отчетем факта, че степента на влияние на ИТМ върху средната продължителност на живота при мъжете е различно от това при жените, трябва да пресметнем

сумата  $[A + B]$ . Събираемото  $B$  отчита ефекта от различната степен на влияние на ИТМ между двата пола във възрастовата група 18 - 24 години, по-точно в съотношение  $w_7^{BMI} * \%BMI_7 : w_1^{BMI} * \%BMI_1$  (приложение, табл. 3).

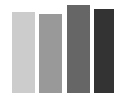
Аналогично, събираемото  $C$  от формула (9) показва колко голяма би била разликата  $life\_exp\_diff_{(18-24)}$ , ако върху нея влияе само факторът **ТЮТЮНО-ПУШЕНЕ** и ако това влияние се осъществява в еднаква степен при мъжете и жените на възраст 18 - 24 години, а именно - тяхната средна аритметична стойност. Ако искаме да отчетем факта, че степента на влияние на тютюнопушенето върху средната продължителност на живота при мъжете е различно от това при жените, трябва да пресметнем сумата  $[C + D]$ . Събираемото  $D$  отчита ефекта от различната степен на влияние на тютюнопушенето между двата пола във възрастовата група 18 - 24 години, по-точно в съотношение  $w_7^{no-sm} * \%No\_smoke_7 : w_1^{no-sm} * \%No\_smoke_1$  (приложение, табл. 3).

## 5. Резултати

### 5.1. Числени резултати от декомпозицията

Съгласно декомпозиционната формула (8) (или съкратеният ѝ запис (9) сумата на ИТМ ефекта и ефекта от въздържане от тютюнопушене, с отчитане на различните им степени на влияние при мъжете и жените в рамките на една и съща възрастова група, е равна на разликата в средните им продължителности на предстоящия живот на съответната възраст. Числените стойности на декомпозиционните ефекти са дадени в табл. 1. Всяка от тях показва каква би била стойността на разликата в средната продължителност -  $life\_exp\_diff_{(AGE\_GROUP)}$ , ако върху нея влияе само и единствено съответният фактор, със или без отчитане на различните степени на влиянието му между двата пола.

Анализът показва как би се променила разликата в средните продължителности на живота между мъжете и жените по възрастови групи, но не прави прогноза кои от тези групи при дадени обстоятелства имат шанс да живеят по-дълго. Във възможностите на анализа е да покаже колко трябва да нарасне продължителността на живота в определена възрастово-полова група, за да стане тя ефективна (при запазване на дяловете на непушачите и лица с ИТМ в норма), но това не е предмет на настоящия анализ. При интерпретацията на резултатите трябва да се има предвид, че изследваните два фактора не обясняват цялата вариация в общия брой години живот. Ако освен тези два фактора в анализа се добавят останалите четири „детерминанти на здравето“ според Европейското здравно интервю, обяснената вариация ще нарасне, което би увеличило практическата познавателна стойност на изследването.



### 1. Стойности на декомпозиционните ефекти върху разликата в средната продължителност на предстоящия живот между мъжете и жените в България за 2008 година

Възрастови групи (години)	ИТМ ефект		Ефект тютюнопушене		Разлика в средната продължителност на предстоящия живот
	без отчитане на различните степени на влияние между половете	с отчитане на различните степени на влияние	без отчитане на различните степени на влияние	с отчитане на различните степени на влияние	
18 - 24	2.16	<b>-0.08</b>	4.88	<b>7.11</b>	<b>7.04</b>
25 - 34	2.22	<b>1.82</b>	4.61	<b>5.00</b>	<b>6.83</b>
35 - 44	1.98	<b>3.58</b>	4.57	<b>2.97</b>	<b>6.54</b>
45 - 54	2.43	<b>2.34</b>	3.68	<b>3.76</b>	<b>6.11</b>
55 - 64	1.57	<b>-3.27</b>	3.49	<b>8.33</b>	<b>5.06</b>
65 - 74	0.88	<b>-3.65</b>	2.33	<b>6.86</b>	<b>3.21</b>

Интерпретацията на резултатите може да се извършва по възрастови групи и по видове ефекти. В най-младата възрастова група 18 - 24 години, разликата в средната продължителност на предстоящия живот между мъжете и жените е най-голяма в сравнение с останалите пет възрастови групи - 7.04 години. Ако допуснем, че в тази група върху средната продължителност на живота влияе само процентът на лицата с ИТМ в норма и отчетем различните степени на влияние на този фактор при мъжете и жените и също така приемем, че всички останали фактори (включително тютюнопушенето) са поставени при равни условия, се получава, че разликата в средната продължителност на живота между мъжете и жените би била почти нулева, дори отрицателна, т.е. в полза на мъжете. Освен в първата възрастова група отрицателен ИТМ ефект се наблюдава и при последните две възрастови групи - 55 - 64 и 65 - 74 години. За тях хипотетичната разлика приема стойности между 3 и 4 години в полза на мъжете, което се обяснява с близките стойности на дяловете лица с ИТМ в норма между двата пола, като дори в последната възрастова група делът при мъжете (36%) надвишава този при жените (30%), което е извън общата тенденция при останалите групи, с изключение на най-младата група. При възрастовите групи от 25 до 54 години ИТМ ефектът е положителен, като най-голямата хипотетична стойност на разликата в продължителността на предстоящия живот се наблюдава при групата 35 - 44 години, а именно 3.58 в полза на жените. Това се дължи на почти два пъти по-големия дял на жени с ИТМ в норма (61.2%) в сравнение с този на мъжете (35.6%).

Ефектът от въздържането от тютюнопушене е по-силен от този на ИТМ в първите две възрастови групи като при 18 - 24-годишните обяснява почти цялата разлика, а при 25 - 34-годишните - пет години от 6.83. В междинните две групи ефектът на неупотребата на цигари е по-слаб в сравнение с по-младите

възрасти, което се съгласува с по-високите положителни стойности на ИТМ ефекта за хората от 35 до 54 години. При последните две възрастови групи (от 55 до 74 години) хипотетичната разлика в средната продължителност на живота между мъжете и жените е с около три години по-голяма от реалната ѝ стойност поради наблюдаваните разлики в дяловете на непушачи в полза на жените и същевременно незначимите разлики при дяловете на имащите нормален индекс на телесна маса сред тях.

## 5.2. Графични изображения. Клъстерен анализ

Проведен е двустъпков клъстерен анализ (Two Step Cluster Analysis) по променливите, разположени на силни скали<sup>11</sup> - „общ брой години”, „дял на непушачите” и „дял на лицата с ИТМ в норма”, като за мярка е използвано **евклидовото разстояние**. За по-голяма нагледност на получените резултати са представени графичните изображения на всеки от разгледаните декомпозиционни фактори, разположени по абсцисата  $Ox$ , и резултатната променлива - по ординатата  $Oy$ . На фиг. 1 е показана зависимостта между относителния дял на лицата с ИТМ в норма и общия брой години живот. Обособяват се пет клъстера:

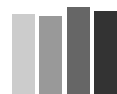
- **Клъстер 1** се състои от двете най-възрастни групи на жените (от 55 до 74 години) и най-възрастната група на мъжете (65 - 74 години). Този клъстер се характеризира с високи стойности на общия брой години живот (над 79 години), високи стойности на дела на непушачите (над 84%) и ниски стойности на дела на лицата с ИТМ в норма (по-малки или равни на 36%).

- **Клъстер 2** се състои от останалите четири групи при жените, а именно жените на възраст от 18 до 54 години. За този изцяло женски клъстер характерни са сравнително високите стойности на общия брой години живот (между 78 и 80 години). Поради големия възрастов диапазон делът на непушачите и този на лицата с ИТМ в норма варират в широки граници. Намаляването на дела на лицата с ИТМ в норма е в права зависимост с увеличаването на възрастта.

- **Клъстер 3** се състои от една-единствена възрастово-полова група, а именно групата на възраст 55 - 64 години при мъжете. Тази група се характеризира с много нисък дял на лицата с ИТМ в норма (30.8%) и сравнително висок спрямо по-младите четири мъжки групи общ брой години живот (близо 76 години).

- **Клъстер 4** също се състои от една-единствена група - най-младата група при мъжете, т.е. на възраст 18 - 24 години. Характеризира се с ниска стойност

<sup>11</sup> Условието за нормално разпределение е налице за всяка от трите променливи. Проверката е осъществена по метода на Колмогоров-Смирнов. Преди анализа данните са сортирани по произволен начин, за да се избегне евентуално повлияване на резултата от определена подредба.

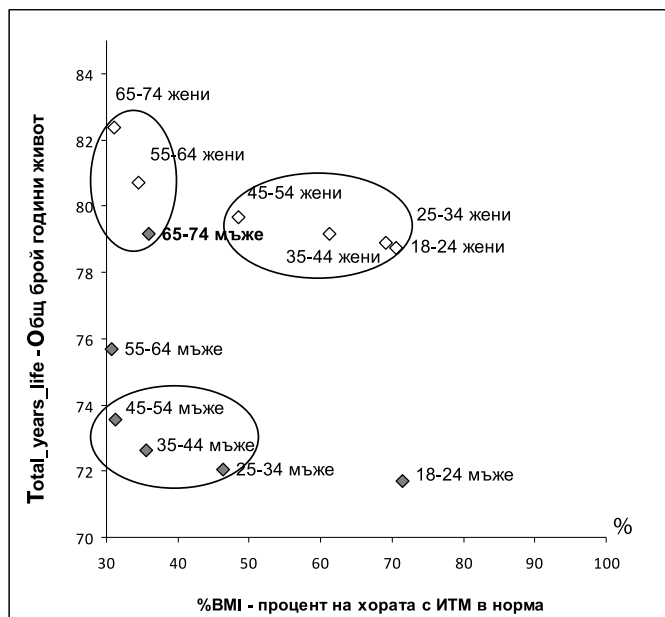


на общия брой години живот (близо 72 години) и същевременно много висока стойност на дела на лицата с ИТМ в норма (71.4%). Забелязва се, че както при жените, така и при мъжете в най-младата възрастова група делът на лицата с ИТМ в норма достига своите максимални стойности.

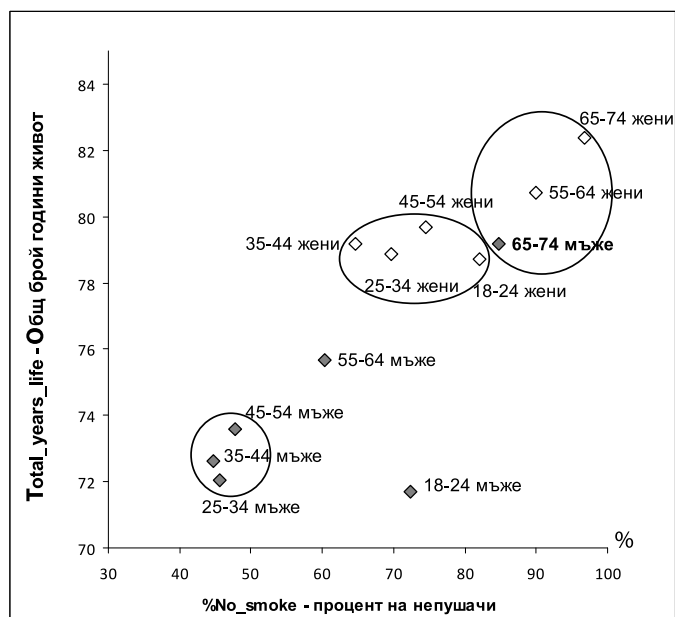
- **Клъстер 5** се състои от междинните три възрастови групи на мъжете, а именно на възраст между 25 и 54 години. Характеризира се със сравнително ниски стойности на общия брой години живот (между 72 и 74 години) и същевременно ниски стойности на дела на лицата с ИТМ в норма. Дяловете на непушачите в този клъстер се движат между 45 и 48% и представляват най-ниските стойности сред разгледаните 12 възрастово-полови групи.

Основният извод от фиг. 1 е, че три от възрастовите групи при мъжете, а именно групите - 18 - 24, 55 - 64 и 65 - 74 години, притежават различни характеристики от останалите групи при мъжете, като последната група дори попада в Клъстер 1, където се намират женските групи с най-висока продължителност на живота. Възрастовите групи при жените се разделят на две - по-младите, сред които непушачите, са по-малко, а имащите ИТМ в норма са повече в сравнение с най-възрастните женски групи, „достигащи“ максимална продължителност на живота. Същата клъстеризация е представена на фиг. 2, където на абсцисата стои процентът на непушачите, а по ординатата - общият брой години живот.

**Фиг. 1. Относителен дял на лицата с ИТМ в норма и общ брой години живот по възрастово-полови групи за 2008 година**



**Фиг. 2. Относителен дял на непушачите и общ брой години живот по възрасто-полови групи за 2008 година**



## 6. Заключение

Освен за сравнение на ефективността на еднородни единици в контекста на производителния процес детерминистичният метод DEA може да се използва в демографската статистика в ролята на декомпозиционна техника. Възрасто-половите групи са обектите на сравнение, декомпозиционните фактори се явяват ресурси, а продължителността на живота е разгледана като техен продукт. DEA пресмята сравнителните тегла на ресурсите, които показват степента на влияние на всеки от тях върху крайния продукт. С помощта на зависимостите между ресурсите и продукта, които характеризират използвания DEA модел и формулата на Kitagawa (1955), се извежда декомпозиционна връзка. В резултат се изчислява ефектът от всеки фактор върху разликата в средната продължителност на предстоящия живот между мъжете и жените. В настоящото изследване разгледаните фактори са два - наличието на ИТМ в норма и въздържането от тютюнопушене.

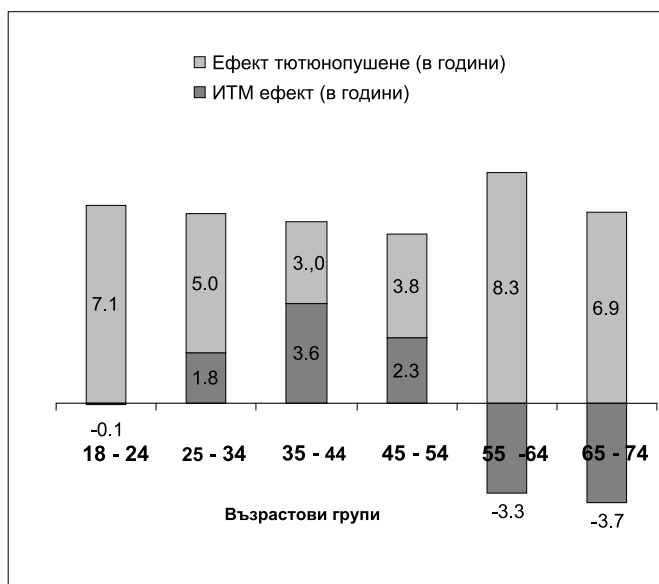
Жените от всички разгледани възрастови групи живеят по-дълго от мъжете и в резултат на настоящия анализ това се **обяснява** най-вече с по-малките дялове на пушачи сред тях. Под „обяснява” не се има предвид каквато и да е причинно-следствена връзка (Kitagawa (1955), pp. 1183-1184), а именно,





че при по-големи разлики в дяловете на пушачите между мъжете и жените в отделните възрастови групи, се наблюдава по-голяма разлика в средните продължителности на предстоящия им живот. ИТМ също има ефект върху предстоящия брой години живот, но степента му на влияние е средно около два пъти по-слаба от тази на тютюнопушенето, като съотношението между тях варира между 1:3 и 2:3, което се доказва от резултатите в приложението, табл. 3. Ако изолираме ефекта от неупотребата на цигари и приемем равни други условия, разликата в средната продължителност на живота се обръща в полза на мъжете при последните две възрастови групи - 55 - 64 и 65 - 74 години, и е почти нулева за групата 18 - 24 години. Причината за това е разположението на тези три групи извън Клъстер 5, който се състои от останалите три възрастови групи при мъжете. При възрастите от 25 до 54 години и двата ефекта са положителни, което означава, че очакваната продължителност на живота при жените е по-голяма както при отчитане само на ИТМ ефекта, така и при отчитане единствено на ефекта от тютюнопушенето. Именно тези три междинни възрастови групи от 25 до 54 години диференцират мъжете и жените и ги обособяват в различни клъстери. Следователно представеният клъстерен анализ е добро средство за онагледяване на резултатите от декомпозицията, осъществена с DEA, както и за тяхната интерпретация. Резултатите от декомпозиционния анализ по двата ефекта са представени на фиг. 3.

**Фиг. 3. Повъзrastова декомпозиция на разликата в средната продължителност на предстоящия живот между мъжете и жените в България за 2008 година**



**ЦИТИРАНА ЛИТЕРАТУРА:**

**Жекова, С.** (2009). Демографска статистика, Икономически университет - Варна.

**Жекова, С.** (2011). Пространствена асиметрия във възрастовата смъртност на населението в България, Известия (Икономически университет - Варна), 1/2011, с. 46 - 60.

**Калоянов, Т.** (2011). Декомпозиция на промените в средната продължителност на предстоящия живот, Население, 1 - 2/2011, с. 3 - 18.

**НСИ**, таблица „Смъртност и средна продължителност на предстоящия живот на населението по пол и местоживеене през периода 2008 - 2010 година”.

**НСИ**, Методология на извадката. Европейско здравно интервю, октомври - ноември 2008.

**Русев, Б.** (1969). Смъртност и средна продължителност на предстоящия живот, Статистика, кн. 2, с. 18 - 35.

**Русев, Б.** (1977). Измерване влиянието на повъзрастовата смъртност върху средната продължителност на живота, Икономическа мисъл, кн. 1, с. 70 - 79.

**Andreev, E. M., Shkolnikov V. M., Begun A. Z.** (2002). Algorithm for decomposition of differences between aggregate demographic measures and its application to life expectancies. healthy life expectancies. parity-progression ratios and total fertility rates, Demographic Research, Vol. 7, Article 14, pp. 499-522.

**Arias, E.** (2003). National Vital Statistics Reports. volume 54, number 14 - United States, Life Tables.

**Arriaga, E. E.** (1982). A Note on the Use of Temporary Life Expectancies for Analyzing Changes and Differentials of Mortality, WHO, Mortality in South and East Asia: A Review of Changing Trends and Patterns, Manila 1980. Geneva: World Health Organization. pp. 559-562.

**Arriaga, E. E.** (1984). Measuring and Explaining the Change in Life Expectancies. Demography, Vol. 21, No. 1, pp. 83-96.

**Castro, M. C.** (2001). Changes in Mortality and Life Expectancy: Some Methodological Issues, Mathematical Population Studies, Vol. 9, pp. 181-208.

**Centers for Disease Control and Prevention (CDC)**, Body Mass Index: Considerations for practitioners.

**Cook, W. D., Zhu J.** (2008). Data Envelopment Analysis, Modeling Operational Processes and Measuring Productivity.

**Cooper, W. W., Seiford L. M., Tone K.** (2007). Data Envelopment Analysis, A Comprehensive Text with Models. Applications, References and DEA-Solver Software, Springer.

**Das Gupta, P.** (1978). A General Method of Decomposing a Difference between Two Rates into Several Components, Demography, Vol. 15, No. 1, pp. 99-112.



**European Health Interview Survey (EHIS)** (2008), implemented and managed by Eurostat.

**Keyfitz, N.** (1977). *Applied Mathematical Demography*, pp. 62-74, Wiley: New York and London.

**Kitagawa, E. M.** (1955). Components of a Difference between Two Rates, *Journal of American Statistical Association*, Vol. 50, No. 272, pp. 1168-1194.

**Pollard, J. H.** (1982). The Expectation of Life and Its Relationship to Mortality, *Journal of the Institute of Actuaries* 109, pp. 225-240.

**Pollard, J. H.** (1988). On the Decomposition of Changes in Expectation of Life and Differentials of Life Expectancy, *Demography*, Vol. 25, No. 2, pp. 265-276.

**Pressat, R.** (1985). The significance of variations in mortality by age on differences in life expectancy, *Population* Vol. 4-5, pp. 762-770.

**Preston, S., Beltran-Sanchez, H.** (2007). A New Method for Attributing Changes in life expectancy to Various Causes of Death, with Application to the United States, PSC Working Paper Series PSC 07-01.

**Retherford, R. D., Cho L. J.** (1973). Comparative Analysis of Recent Fertility Rates in East Asia, *International Union for Scientific Study of Population, Proceedings of 17-th General Conference of the IUSSP*, Vol. 2, pp. 163-181.

**Romo, V. C.** (2003). *Decomposition methods in demography*, Amsterdam Rozenberg Publishers.

**United Nations** (1982). *Levels and Trends of Mortality Since 1950, Study 74*. New York: United Nations, Dept. of International Economic and Social Affairs.

**Vaupel, J. W.** (1986). How Change in Age-Specific Mortality Affects Life Expectancy, *Population Studies*, Vol. 40, pp. 147-157.

**Vaupel, J. W., Romo V. C.** (2002). Decomposing Demographic Change into Direct vs. Compositional Components. *Demographic Research*. Vol. 7. pp. 1-14.

**Vaupel, J. W., Romo, V. C.** (2003). Decomposing Change in Life Expectancy: a Bouquet of Formulas in honor of Nathan Keyfitz's 90-th Birthday, *Demography*, Vol. 40, No. 2, pp. 201-216.

**World Health Organization (WHO)**, Body mass index.

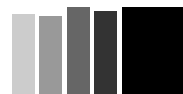
## ПРИЛОЖЕНИЕ

### 1. Относителни дялове на лицата с ИТМ в норма и на непушачите и общ брой години живот по възрасто-полови групи за 2008 година

№ на възрасто-полова група	Възрасто-полови групи	Процент на хората с ИТМ в нормата %BMI	Процент на непушачите %No_smoke	Общ брой години живот total_years_life
1	18 - 24 мъже	71.40	72.30	71.70
2	25 - 34 мъже	46.30	45.60	72.06
3	35 - 44 мъже	35.60	44.70	72.62
4	45 - 54 мъже	31.20	47.80	73.57
5	55 - 64 мъже	30.80	60.30	75.67
6	65 - 74 мъже	36.00	84.70	79.17
7	18 - 24 жени	70.50	82.10	78.73
8	25 - 34 жени	69.10	69.70	78.88
9	35 - 44 жени	61.20	64.70	79.16
10	45 - 54 жени	48.40	74.50	79.68
11	55 - 64 жени	34.40	90.00	80.72
12	65 - 74 жени	31.00	96.80	82.38

### 2. Нива на ефективност и оптимални множители в резултат на DEA по възрасто-полови групи за 2008 година

№ на възрасто-полова група	Възрасто-полови групи	Ниво на ефективност	Оптимален множител на %BMI	Оптимален множител на %No_smoke	Оптимален множител на total_years_life
1	18 - 24 мъже	0.57	0.005	0.009	0.008
2	25 - 34 мъже	0.90	0.007	0.015	0.013
3	35 - 44 мъже	1.00	0.008	0.016	0.014
4	45 - 54 мъже	1.00	0.013	0.013	0.014
5	55 - 64 мъже	0.89	0.011	0.011	0.012
6	65 - 74 мъже	0.71	0.008	0.008	0.009
7	18 - 24 жени	0.58	0.004	0.009	0.007
8	25 - 34 жени	0.65	0.005	0.010	0.008
9	35 - 44 жени	0.72	0.005	0.011	0.009
10	45 - 54 жени	0.70	0.008	0.008	0.009
11	55 - 64 жени	0.70	0.008	0.008	0.009
12	65 - 74 жени	0.69	0.008	0.008	0.008



### 3. Степени на влияние на факторите върху общия брой години живот по възрастово-полови групи за 2008 година

№ на възрастово-полова група	Възрастово-полови групи	$w^{BMI} * \%BMI$ - степен на влияние на ИТМ върху <b>total_years_life</b>	$w^{no\_sm} * \%No\_smoke$ - степен на влияние на тютюнопушенето върху <b>total_years_life</b>
1	18 - 24 мъже	0.32	0.68
2	25 - 34 мъже	0.33	0.67
3	35 - 44 мъже	0.29	0.71
4	45 - 54 мъже	0.40	0.60
5	55 - 64 мъже	0.34	0.66
6	65 - 74 мъже	0.30	0.70
7	18 - 24 жени	0.29	0.71
8	25 - 34 жени	0.32	0.68
9	35 - 44 жени	0.31	0.69
10	45 - 54 жени	0.40	0.60
11	55 - 64 жени	0.28	0.72
12	65 - 74 жени	0.25	0.75

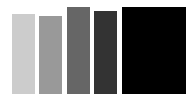
## ВОЗДЕЙСТВИЕ ИНДЕКСА МАССЫ ТЕЛА И КУРЕНИЯ НА РАЗНИЦУ В СРЕДНЕЙ ПРОДОЛЖИТЕЛЬНОСТИ ПРЕДСТОЯЩЕЙ ЖИЗНИ

*Петя Брайнова\**

**РЕЗЮМЕ** В 2008 году, по данным Евростата, разница в ожидаемой продолжительности предстоящей жизни между мужчинами и женщинами в Болгарии в возрасте 18 лет, составляет немного больше семи лет. В данной статье представлена возрастная декомпозиция этой разницы в зависимости от воздействия индекса массы тела и курения в шести возрастных группах (18 - 24, 25 - 34, 35 - 44, 45 - 54, 55 - 64 и 65 - 74 лет). Для этой цели используется Data Envelopment Analysis (DEA), который сравнивает однородные единицы на основе измерения их эффективности подобно соотношению продукции и использованных для ее производства ресурсов. Каждая из возрастных групп является в роли однородной единицы сравнения, при этом рассматриваемых единиц в целом двенадцать - шесть для мужчин и шесть для женщин. Сумма прожитых к данному моменту и последующих с этого момента лет рассматривается как продукт здорового образа жизни, характеризующегося наличием индекса массы тела в норме и воздержания от курения.

---

\* Докторант на Кафедре статистики и эконометрии Университета национального и мирового хозяйства, София; e-mail: petia\_brainova@abv.bg.



## THE BODY MASS INDEX EFFECT AND TOBACCO SMOKING EFFECT ON THE DIFFERENCE IN LIFE EXPECTANCY

*Petya Braynova\**

**SUMMARY** According to the data from Eurostat for 2008, the difference in life expectancy between men and women in Bulgaria, at 18 years of age is a little over 7 years. The current article presents an age-decomposition of the difference in life expectancy by the Body mass index effect and the tobacco smoking effect separately for 6 age groups (18 - 24, 25 - 34, 35 - 44, 45 - 54, 55 - 64 and 65 - 74). For that purpose, Data Envelopment Analysis (DEA) is applied. DEA is a tool for comparison of homogeneous units based on measurements of their efficiency, defined as a ratio of the products and the resources used in the production process. Each of the age groups represents a homogeneous unit. The total number of units is 12 - 6 units for men and 6 units for women. The total amount of the currently lived years and the years thereafter is considered as the product of a healthy lifestyle, characterized by the presence of normal body mass index and restraint from tobacco smoking.

---

\* PhD student at the Statistics and Econometrics Academic Department, University of National and World Economy, Sofia; e-mail: [petia\\_brainova@abv.bg](mailto:petia_brainova@abv.bg).