

---

# Методически подход за определяне на представителна извадка от земеделски стопанства в България

Доц. д-р МИНКА АНАСТАСОВА-ЧОПЕВА

*Институт по аграрна икономика – София*

E-mail: [chopevam@yahoo.com](mailto:chopevam@yahoo.com)

Целта на статията е да представи методологията за определяне на представителна извадка от земеделски стопанства и резултатите от нейното приложение. Поставената цел е обусловена от факта, че голяма част от научните изследвания на земеделските стопанства се базират на информация от извадкови анкетни проучвания. Използван е методът на районирането, съчетан с модела на двустепенната гнездова извадка. При определянето на броя и състава на областите (гнездата) и земеделските стопанства са съблюдавани изискванията, гарантиращи възможността в извадката да попаднат областите и съответно ЗС, които възпроизвеждат в достатъчно висока степен структурата на земеделските стопанства по тяхната специализация.

Изчислени са две дисперсии: едната, наречена междугнездова, която представлява величината на дисперсията във всеки един район на планиране, а втората се отнася до дисперсията вътре в самите области или т.н. вътрешногнездова дисперсия. Това е извършено, за да се постигне необходимото съответствие между структурата на ЗС от генералната съвкупност по производствената специализация и тяхната структура в извадката. Допълнително са използвани още два показателя за доказване на представителността на извадката. Първият отразява степента на съответствие между средния брой земеделски стопанства в едно гнездо (област) от извадката и средния брой земеделски стопанства в една област от генералната съвкупност. Вторият показател измерва степента на съответствие между вариацията в избраните гнезда от извадката и общата вариация в генералната съвкупност.

Имайки предвид получените стойности на междугнездовата и вътрешногнездовата дисперсия при избора на областите, които ще попаднат в извадката, сме следвали следното правило. Първо, от всеки район на планиране е избрана тази област, която най-близко се доближава до средната за съответния район на планиране производствена структура на земеделските стопанства. Второ, допълнително са избрани тези области, които имат най-силно отклонение от производствената структура, изразяваща средната за района структура на земеделските стопанства. По този начин се гарантира, че в извадката ще попаднат области, които в достатъчна степен ще обхванат производственото разнообразие в съответния район и ще възпроизведат структурата на земеделските стопанства.

В обхвата на извадката са включени стопанства от 20 области в страната. Областите, от които липсват земеделски стопанства в извадката, са такива, чиято структура по специализация е много близка до структурата на съседни на тях области.

Стойността на максимално допустимата грешка при гаранционен множител ( $t = 1,96$ ) е равна на 4,74%. Следователно може да се твърди с 95% гаранционна вероятност, че действителните стойности на оценяваните параметри от анкетната информация ще се движат в границите на доверителния интервал –  $\pm 4,27\%$ .

Един от основните методически подходи на научните изследвания в областта на земеделските стопанства е анкетният метод. За целта обикновено се провежда анкетно проучване на част от земеделските производители (извадка) по въпроси, които са предмет на съответното научно изследване. Понеже такъв вид изследвания не са изчерпателни, а извадкови, получените резултати и на-

правените изводи са напълно валидни само за извадката, а що се отнася до всички земеделски стопанства, те са валидни с определена гаранционна вероятност. Тази вероятност се определя от степента на представителност на извадката. Колкото тя е по-голяма по обем и възпроизвежда структурата на генералната съвкупност по даден признак или признаци, толкова е по-висока гаранционната вероят-

ност. Това налага изготвянето на представителна за страната извадка от земеделски стопанства, които могат да бъдат обект на анкетиране, и получените резултати да се екстраполират за генералната съвкупност с определена степен на гаранционна вероятност. Получените резултати от приложението на методологията за изготвяне на представителна извадка от ЗС са използвани при разработването на научен проект „Влияние на инвестиционната подкрепа върху икономическата жизнеспособност на земеделските стопанства”. Проектът се осъществява от колектив на ИАИ, с ръководител проф. д-р Д. Николов, през периода 2015–2016 г.

**Целта** на статията е да представи методологията за определяне на представителна извадка от земеделски стопанства и резултатите от нейното приложение.

## Методология

За постигане на набелязаната цел е използван методът на районирането, комбиниран с модела на двустепенната гнездова извадка. Като критерий за райониране е използван териториалният признак. На първата степен от формирането на извадката са избрани в качеството на гнезда 20 области от 6-те района на планиране. На втората степен, спазвайки принципа на пропорционалността, са избрани необходимият брой земеделски стопанства от всяко гнездо (Гатев, К., 1986; Kish, L., 1965; Loga, L. Bohn, 1996).

При определянето на броя и състава на областите (гнездата), и земеделските стопанства са съблюдавани изискванията, гарантиращи възможността в извадката да попаднат областите и съответно ЗС, които възпроизвеждат в достатъчно висока степен структурата на земеделските стопанства по тяхната специализация. Съгласно възприетата типология на ЗС от МЗХ (Преброяване на земеделските стопанства в България през 2010 г.), земеделските стопанства във всяка област са обособени според производствената им специализация в следните основни групи: полски култури; зеленчуци, цветя и гъби; лозя;

трайни насаждения; всички видове говеда; овце, кози и други тревопасни; свине, птици и зайци; стопанства от смесен тип, независимо дали преобладава отглеждането на животни или земеделски култури. За да се постигне необходимото съответствие между структурата на производствената специализация на ЗС от генералната съвкупност и тази в извадката, са изчислени две дисперсии: едната, наречена междугнездова, която представлява величината на дисперсията във всеки един район на планиране, а втората се отнася до дисперсията вътре в самите области или т.н. вътрешногнездова дисперсия.

Определянето на средната грешка става по метода на конструирането и се извършва съгласно следната формула (Съйкова, И., Б. Чакалов, 1977):

$$1) \mu_{0r}^2 = (\sum N_h^2 \mu_{0h}^2) / (\sum N_h)^2, \text{ където:}$$

$N_h$  е общият брой на земеделските стопанства в район  $h$  ( $h = 1, 2 \dots 6$ );

$\mu_{0h}^2$  представлява районната стохастична грешка и се определя посредством следната формула:

$$2) \mu_{0h}^2 = (\sigma_{0rh}^2 / m_h) \times (1 - m_h / M_h) + (\sigma_{bh}^2 / m_h * M_h) \times (1 - \check{n}_h / \hat{n}), \text{ където:}$$

$\sigma_{0rh}^2$  е междугнездовата дисперсия;

$\sigma_{bh}^2$  е вътрешногнездовата дисперсия;

$m_h$  е броят на гнездата (областите) за наблюдение в район  $h$ ;

$M_h$  е общият брой гнезда (области) в район  $h$ ;

$\check{n}_h$  е средният брой единици (земеделски стопанства) за наблюдение в едно гнездо в район  $h$ ;

$\hat{n}$  е средният брой единици (земеделски стопанства) в едно гнездо в район  $h$ .

Междугнездовата и вътрешногнездовата дисперсия са модифицирани форми на общата формула, по която се изчислява дисперсията при обикновената извадка:

$$3) \sigma^2 = \sum (X_{idh} - X_{dh})^2 / (n_{dh} - 1), \text{ където:}$$

с индекс  $i$  се идентифицира типът на земеделското стопанство ( $i = 1, \dots, 8$ );

$X_{idh}$  означава броят (или относителният дял) на ЗС в област  $d$  от район  $h$ ;

$X_{dh}$  е средният брой на ЗС в област  $d$  от район  $h$ ;

$n_{dh}$  е броят на различните групи ЗС.

След определяне на средната стохастична грешка се изчислява т.н. максимално допустима грешка на оценката, при което се взема предвид доверителната вероятност и съответният гаранционен множител  $t$ . Най-често използваните стойности са 95% гаранционна вероятност на получените оценъчни резултати и  $t = 1,96$ . Формулата, по която се изчислява максимално допустима грешка, има следния вид:

$$4) \Delta_{\theta} = \pm t \times \mu_{\theta r}$$

Общата формула, по която се определя обемът на извадката в елементарни единици (в случая земеделски стопанства) при двустепенната гнездова извадка, е следната:

5)  $n_{\theta} = (t^2 \times \sigma_{\theta}^2) / (\Delta_{\theta}^2)(1 + \delta \times (\bar{n}-1))$ , където отделните символи имат смисловото значение, отбелязано по-горе.

Добре е единиците на извадката (ЗС) да се изберат по възможност на случаен принцип, за да се гарантира в максимална степен нейната стохастичност.

След като е определен списъкът с областите, които попадат в извадката, допълнително са използвани още два показателя за доказване целесъобразността на техния избор.

Първият показател се отнася до степента на приближение на средния брой земеделски стопанства в едно гнездо (област) от извадката до средния брой земеделски стопанства в една област от генералната съвкупност. Количествено този показател се изразява със следното съотношение:

$$a) P_1 = n_{изв./обл.} / n_{ген.с./обл.}$$

Вторият критерий измерва степента на доближаване на вариацията в избраните гнезда от извадката до общата вариация в генералната съвкупност. Формулата, по която се изчислява обхватът на вариацията, е следната:

$$б) P_2 = VAR_{изв.} / VAR_{ген.с\&v\&k.}$$

И за двата допълнителни показателя важи правилото: стойностите на  $P_1$  и  $P_2$  да са достатъчно близки до 1.

## Резултати и изводи

На първия етап получените резултати показват, че в извадката са включени 20 области от 6-те райони на планиране. Това са областите: Монтана, Ловеч, Плевен, Силистра, Разград, В. Търново, Търговище, Варна, Добрич, Бургас, Сливен, Ямбол, Кърджали, Смолян, Пазарджик, Пловдив, Благоевград, Кюстендил, Перник и София – област. Както се вижда, броят на избраните области от всеки район на планиране не е пропорционален на техния общ брой в съответния район. Причината е в различната степен на разсейване между областите при отделните райони на планиране и вътре в самите области (табл. 1).

Както се вижда, в Югозападния и Южния Централен район на планиране стойностите на междугнездовата дисперсия, в сравнение с останалите райони, са по-високи. Това на свой ред обуславя необходимостта от по-голям брой единици (земеделски стопанства) от тези райони да попаднат в извадката.

**Таблица 1.** Стойности на междугнездовата и вътрешногнездовата дисперсия

**Table 1.** Values of inter-cluster and intra-cluster dispersion

| Райони на планиране /<br>Planning regions | Общ брой ЗС /<br>Number of farms | Междугнездова дисперсия /<br>Inter-cluster dispersion | Вътрешногнездова<br>дисперсия / Intra-cluster<br>dispersion |
|---|----------------------------------|---|---|
| СЗР / NWR                                 | 51250                            | 27,26   | 15,87   |
| СЦР / NCR                                 | 43281                            | 23,61   | 14,90   |
| СИР / NER                                 | 43657                            | 34,14   | 161,53  |
| ЮИР / YUER                                | 56750                            | 27,87   | 143,44  |
| ЮЦР / YUCR                                | 109035                           | 46,70   | 145,56  |
| ЮЗР / YUWR                                | 64938                            | 50,23   | 159,57  |

Източник: Преброяване на ЗС, 2010 г., (МЗХ) и собствени изчисления.

Source: Census of farms, 2010, (MAF) and own calculations.

В Южния Централен район и броят на земеделските стопанства е най-голям, за разлика от Югозападния район, което предопределя (заедно с високите стойности на междугнездовата) най-голям брой ЗС да попаднат в извадката. Отделно, интерес за разглеждане представлява Североизточният район на планиране. При него вътрешногнездовата дисперсия е много висока, но поради по-големия среден размер на едно ЗС в района, в сравнение с останалите, относително по-малък е общият брой на земеделските стопанства в него.

Освен това, стойността на междугнездовата дисперсия е на средно равнище. Затова необходимият брой ЗС в извадката не трябва да бъде по-голям от този в останалите райони на планиране. В Северозападния район съществуват много на брой, но малки по размер стопанства. Затова, въпреки ниските стойности на разсейване, в извадката е необходимо да попаднат относително по-голям брой ЗС.

Информация за вътрешногнездовата дисперсия във всяка област може да се проследи в табл. 2.

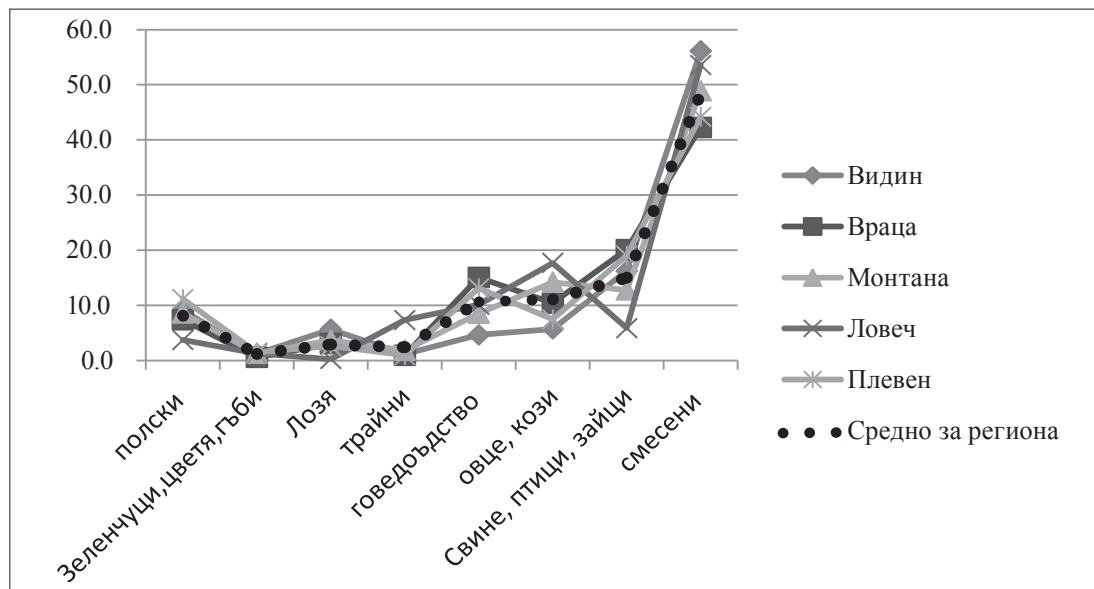
**Таблица 2.** Стойности на вътрешногнездовата дисперсия по области  
**Table 2.** Values of intra-cluster dispersion by districts

| Област / District                | Общ брой ЗС /<br>Number of farms | Икономически размер<br>на 1 ЗС (СП-хил. евро) /<br>Economic size of 1 farm<br>(SP-thousand euro) | Вътрешногнездова<br>дисперсия в областите /<br>Intra-cluster dispersion in<br>districts |
|----------------------------------|----------------------------------|--|---|
| Благоевград / Blagoevgrad        | 38467                            | 2,03   | 122,29  |
| Бургас / Burgas                  | 18662                            | 7,11   | 93,40   |
| Варна / Varna                    | 11758                            | 7,95   | 175,44  |
| Видин / Vidin                    | 8362                             | 4,87   | 18,28   |
| Враца / Vratsa                   | 10216                            | 8,96   | 13,86   |
| В. Търново / Veliko Tarnovo      | 9344                             | 13,87  | 14,39   |
| Габрово / Gabrovo                | 3223                             | 5,84   | 14,55   |
| Добрич / Dobrich                 | 10823                            | 17,06  | 158,46  |
| Кърджали / Kardzhali             | 22833                            | 2,52   | 243,34  |
| Кюстендил / Kyustendil           | 7514                             | 2,83   | 197,08  |
| Ловеч / Lovech                   | 10118                            | 4,79   | 17,49   |
| Монтана / Montana                | 9653                             | 8,21   | 15,46   |
| Пазарджик / Pazardzhik           | 21404                            | 3,88   | 75,16   |
| Перник / Pernik                  | 3594                             | 5,78   | 236,24  |
| Плевен / Pleven                  | 12924                            | 11,65  | 14,26   |
| Пловдив / Plovdiv                | 30445                            | 6,89   | 55,41   |
| Разград / Razgrad                | 9301                             | 10,15  | 14,09   |
| Русе / Ruse                      | 7945                             | 13,65  | 14,58   |
| Силистра / Silistra              | 13468                            | 7,38   | 16,88   |
| Сливен / Sliven                  | 14681                            | 5,52   | 99,00   |
| Смолян / Smolyan                 | 16348                            | 1,51   | 238,12  |
| София - град / Sofia - town      | 1545                             | 7,20   | 96,51   |
| София - област / Sofia -district | 13818                            | 4,28   | 145,75  |
| Ст. Загора / Stara Zagora        | 15696                            | 9,77   | 120,47  |
| Търговище / Targovishte          | 10425                            | 7,14   | 145,72  |
| Хасково / Haskovo                | 18420                            | 6,36   | 115,75  |
| Шумен / Shumen                   | 10740                            | 8,82   | 166,52  |
| Ямбол / Yambol                   | 7906                             | 13,27  | 260,88  |

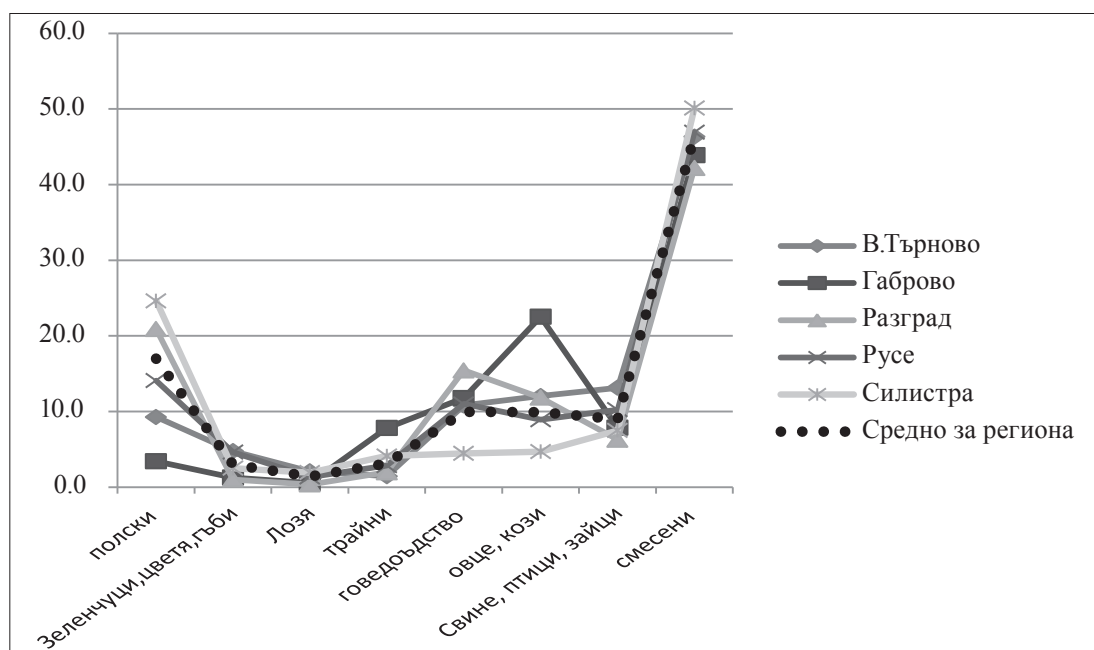
Източник: Преброяване на ЗС, 2010 г., (МЗХ) и собствени изчисления.  
Source: Census of farms, 2010, (MAF) and own calculations.

При определянето както на броя и състава на областите в извадката от съответните райони на планиране, така и на броя на земеделските стопанства от всяка област в извадката, освен показателите за степен на разсейване, е взета предвид и структурата на ЗС във всяка област по производствена специализация. Това се налага поради следното обстоя-

телство: Понякога близки по стойност междугнездова или вътрешногнездова дисперсии в някои райони, съответно области, могат да имат съвсем различни структури на ЗС от гледна точка на тяхната производствена специализация. Например в Северозападен район Плевен и Монтана са със сходни стойности на вътрешногнездовата дисперсия (съот-



Фиг. 1. Разпределение на ЗС в СЗР по производствена специализация (%)  
 Fig. 1. Distribution of farms in NWR by specialization (%)

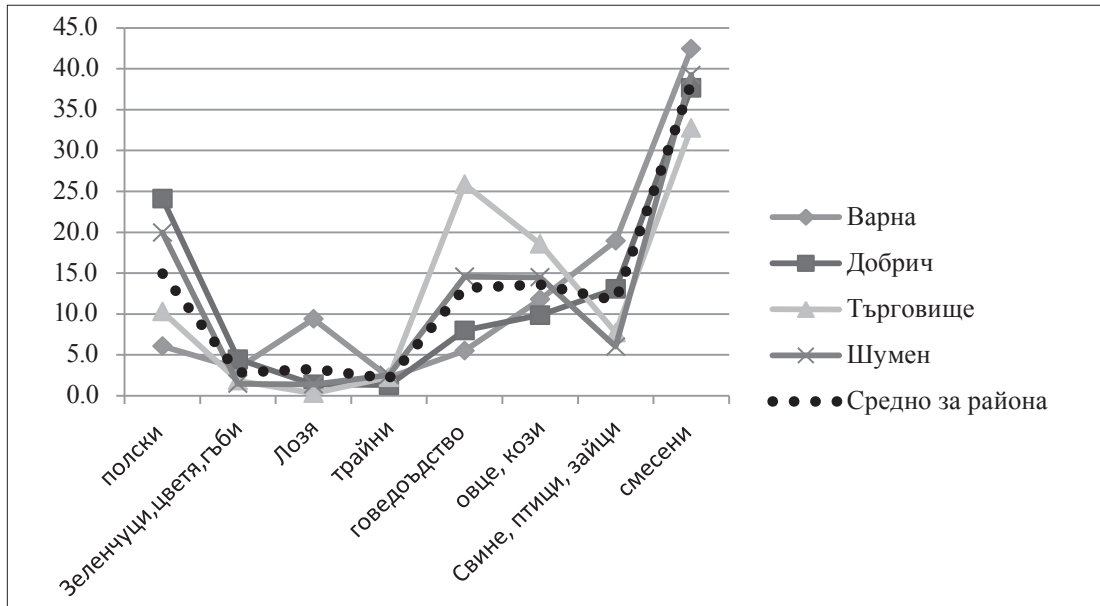


Фиг. 2. Разпределение на ЗС в СЦР по производствена специализация (%)  
 Fig. 2. Distribution of farms in NCR by specialization (%)

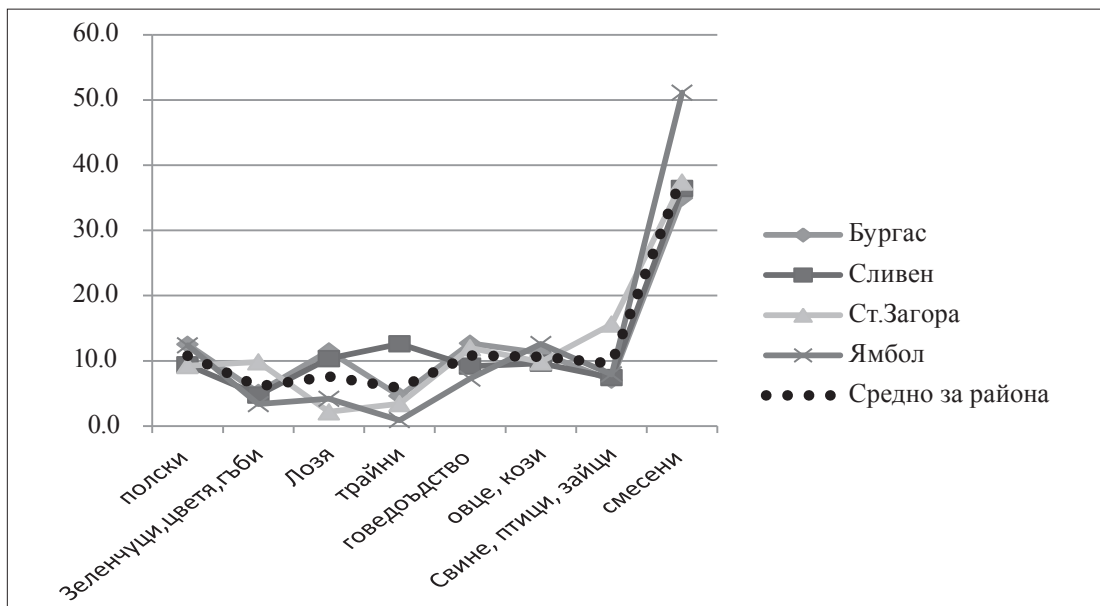
ветно 14,26 и 15,46), но структурите им не са толкова близки. Това може да се проследи на фиг. 1. Относителният дял на стопанствата, които отглеждат овце и кози в област Монтана, е почти два пъти по-висок от този в Плевенска област.

Във всеки един район могат да се посочат подобни примери, което се вижда на графиките с номера от 2 до 6, включително.

Имайки предвид степента на разсейване между отделните области във всеки район на планиране от гледна точка на производствената структура на земеделските стопанства, сме следвали следното правило при избора на областите, които ще попаднат в извадката. Първо, от всеки район на планиране е избрана тази област, която най-близко се доближава до средната за съответния район на плани-

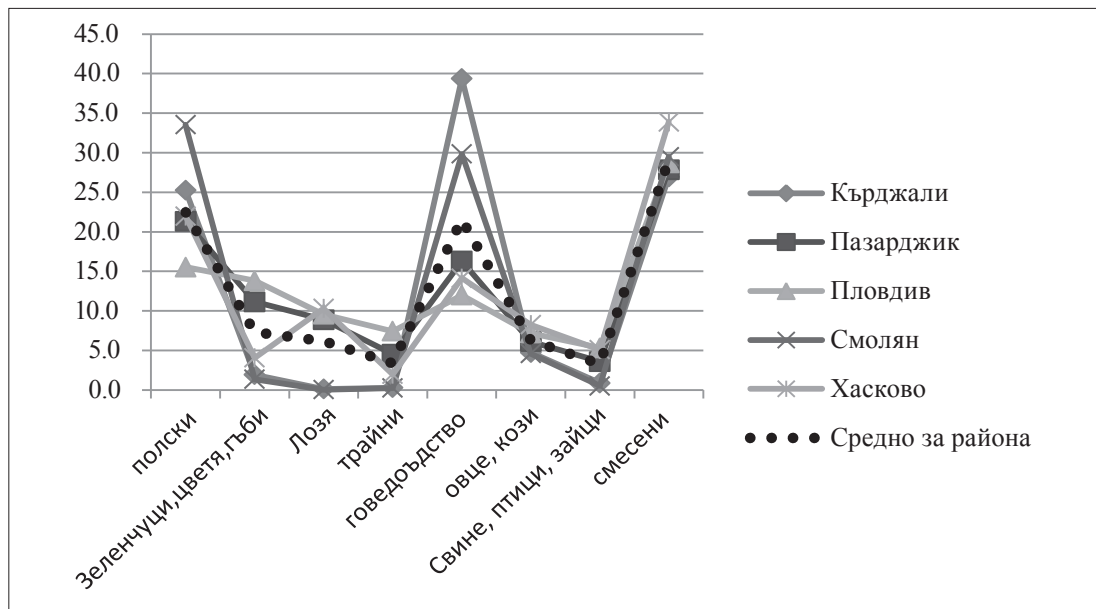


Фиг. 3. Разпределение на ЗС в СИР по производствена специализация (%)  
 Fig. 3. Distribution of farms in NER by specialization (%)

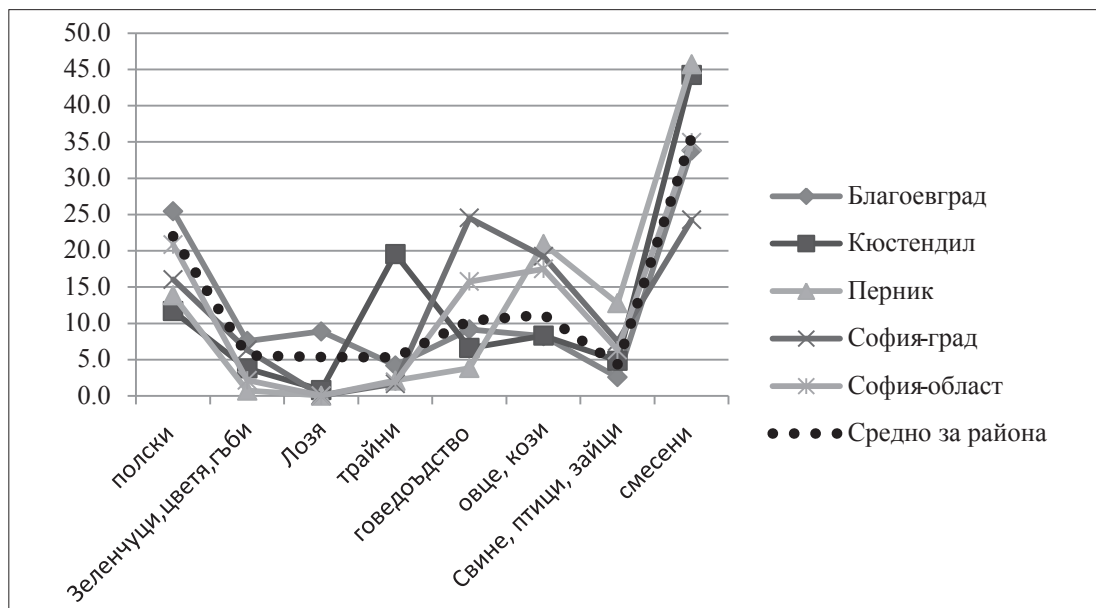


Фиг. 4. Разпределение на ЗС в ЮИР по производствена специализация (%)  
 Fig. 4. Distribution of farms in YUER by specialization (%)





Фиг. 5. Разпределение на ЗС в ЮЦР по производствена специализация (%)  
 Fig. 5. Distribution of farms in YUCR by specialization (%)



Фиг. 6. Разпределение на ЗС в ЮЗР по производствена специализация (%)  
 Fig. 6. Distribution of farms in YUWR by specialization (%)

ране производствена структура на земеделските стопанства. Второ, допълнително са избрани тези области, които имат най-силно отклонение от производствената структура, изразяваща средната за района структура на земеделските стопанства. По този начин се гарантира, че в извадката ще попаднат об-

ласти, които в достатъчна степен ще обхванат производственото разнообразие в съответния район и ще възпроизведат структурата на земеделските стопанства.

Освен това се има предвид числеността и размерът на земеделските стопанства в отделните райони. Например област Габро-

во се отклонява от средната линия на производствената структура в СЦР, но като се вземе предвид, че абсолютният брой на ЗС в областта е най-малък (3223) и делът на ЗС в областта е едва 7,4%, то отпада целесъобразността от нейното включване в извадката. Благоевград като типична земеделска област е включена в извадката, като производствената структура в нея е близка до средната за ЮЗР. Освен това, поради наличието на голям брой и сравнително малки по размер земеделски стопанства в област Благоев-

град, техният брой в извадката е сравнително голям. За всяка една област по подобен начин са разгледани характеристиките ѝ от гледна точка на степен на разсейване, брой и размер на ЗС и е определено кои области, и съответно с колко на брой земеделски стопанства участват в извадката. В резултат на приложената методика за определяне на обема и структурата на извадковата съвкупност се получи следният списък (табл. 3) с областите и броя на земеделските стопанства към тях.

**Таблица 3.** Брой на ЗС в извадката по области и райони на планиране  
**Table 3.** Number of farms in the sample by districts and planning regions

|                                      | Общ брой ЗС /<br>Total number of<br>farms | В т. число / Including:             |   |                    |
|--------------------------------------|---|-------------------------------------|---|--------------------|
|                                      |   | Растениевъдни /<br>Plant production | Животноводни /<br>Livestock<br>production | Смесени /<br>Mixed |
| <b>I. СЗР / NWR</b>                  | <b>42</b>                                 | <b>7</b>                            | <b>15</b>                                 | <b>20</b>          |
| 1. Монтана / Montana                 | 12  | 2                                   | 4   | 6                  |
| 2. Ловеч / Lovech                    | 13  | 2                                   | 4   | 7                  |
| 3. Плевен / Pleven                   | 17  | 3                                   | 7   | 7                  |
| <b>II. СЦР / NCR</b>                 | <b>35</b>                                 | <b>10</b>                           | <b>9</b>                                  | <b>16</b>          |
| 1. Силистра / Silista                | 15  | 5                                   | 2   | 8                  |
| 2. Разград / Razgrad                 | 10  | 3                                   | 3   | 4                  |
| 3. В. Търново / V. Tarnovo           | 10  | 2                                   | 4   | 4                  |
| <b>III. СИР / NER</b>                | <b>35</b>                                 | <b>10</b>                           | <b>13</b>                                 | <b>12</b>          |
| 1. Търговище / Targovishte           | 11  | 3                                   | 5   | 3                  |
| 2. Варна / Varna                     | 12  | 3                                   | 4   | 5                  |
| 3. Добрич / Dobrich                  | 12  | 4                                   | 4   | 4                  |
| <b>IV. ЮИР / YUER</b>                | <b>46</b>                                 | <b>15</b>                           | <b>13</b>                                 | <b>18</b>          |
| 1. Бургас / Burgas                   | 20  | 7                                   | 6   | 7                  |
| 2. Сливен / Sliven                   | 16  | 6                                   | 4   | 6                  |
| 3. Ямбол / Yambol                    | 10  | 2                                   | 3   | 5                  |
| <b>V. ЮЦР / YUCR</b>                 | <b>89</b>                                 | <b>36</b>                           | <b>29</b>                                 | <b>24</b>          |
| 1. Кърджали / Kardzhali              | 24  | 7                                   | 11  | 6                  |
| 2. Смолян / Smolyan                  | 16  | 5                                   | 6   | 5                  |
| 3. Пазарджик / Pazardzhik            | 23  | 11                                  | 6   | 6                  |
| 4. Пловдив / Plovdiv                 | 26  | 13                                  | 6   | 7                  |
| <b>VI. ЮЗР / YUWR</b>                | <b>53</b>                                 | <b>20</b>                           | <b>15</b>                                 | <b>18</b>          |
| 1. Благоевград / Blagoevgrad         | 22  | 10                                  | 5   | 7                  |
| 2. Кюстендил / Kyustendil            | 11  | 5                                   | 2   | 4                  |
| 3. Перник / Pernik                   | 8   | 2                                   | 3   | 3                  |
| 4. София – област / Sofia – district | 12  | 3                                   | 5   | 4                  |
| <b>Общо / Total</b>                  | <b>300</b>                                | <b>98</b>                           | <b>94</b>                                 | <b>108</b>         |

Източник: Собствени изчисления.  
Source: Own calculations.



При така изготвения обем и структура на извадката се получава, че средната грешка ( $\mu_{\theta}^2$ ) има стойност, равна на 4,74%. Съответно максимално допустимата грешка ( $\Delta_{\theta}$ ) при гаранционен множител ( $t = 1,96$ ) ще бъде равна на  $\sqrt{4,74 \times 1,96} = 4,27\%$ . Тогава може да се твърди с 95% гаранционна вероятност, че действителните стойности на оценяваните параметри от анкетната информация ще се движат в границите на доверителния интервал  $\pm 4,27\%$ .

За цялата извадка броят на смесените земеделски стопанства в незначителна степен надхвърля този на растениевъдните и животновъдните стопанства. Това разпределение на ЗС съответства на тяхното разпределение в генералната съвкупност. Отклонение от горепосочената тенденция прави Южният Централен район в извадката, в който преобладават растениевъдните стопанства.

Получените резултати от използваните допълнителни два критерия за представителност на извадката са следните:

- Отношението на средния брой на земеделските стопанства в една област от извадката към средния брой на земеделските стопанства в една област в цялата генерална съвкупност е равно на 1,15.

- Отношението на вариацията между броя на земеделските стопанства в областите от извадката и вариацията между броя на земеделските стопанства между всички области в страната възлиза едва на 1,05.

Както се вижда и двата показателя са близки до 1, особено вторият, свързан със степента на обхват на общата вариация от тази в извадката. Тези стойности дават допълнително основание да се счита, че така определените модел, състав и структура на извадката са достатъчно представителни и възпроизвеждат производствената структура на земеделските стопанства в генералната съвкупност.

За да се осигури доказаната представителност на извадката, от особено значение е да се приложи методът на собствено-слу-

чайния подбор при избор на конкретните земеделски стопанства, които ще бъдат изследвани.

## Заклучение

При изготвянето на обема и структурата на извадката от земеделски стопанства са съблюдавани основните правила за постигане на необходимата степен на нейната представителност. В обхвата на извадката са включени стопанства от 20 области в страната. Областите, от които липсват земеделски стопанства в извадката, са такива, чиято структура по специализация е много близка до структурата на съседни на тях области. Затова евентуалното включване на стопанства от липсващите в горния списък области няма да допринесе за намаляване на стохастичната грешка, а само би оскъпило провеждането на съответното анкетно проучване. По този начин е гарантирано, че структурата на земеделските стопанства по производствена специализация от извадката в максимална степен възпроизвеждат структурата на ЗС в генералната съвкупност. С най-голям брой ЗС в извадката са областите Пловдив, Кърджали и Благоевград. Този резултат съответства на разпределението на стопанствата в генералната съвкупност.

## ЛИТЕРАТУРА

**Гатев, К.** 1986. Обща теория на статистиката. София.

**Съйкова, И., Б. Чакалов.** 1977. Методология и методика на социологическите изследвания. София.

**Kish, L.** 1965. Survey Sampling. Journal of the Royal Statistical Society. Series A (General) Vol. 132, № 2, pp. 272-274

**Lora, L. Bohn.** 1996. A review of nonparametric ranked-set sampling Methodology. Communications in Statistics – Theory and Methods, Vol. 25, Issue 11.

Преброяване на земеделските стопанства в България. 2010. МЗХ.

## Methodological Approach for Determining Representative Sample of Bulgarian Farms

MINKA CHOPEVA

*Institute Agricultural Economics – Sofia*

(Summary)

The aim is to present the methodology for determining a representative sample of farms and the results of its application. The goal is conditioned by the fact that much of the farm research is based on information from sample surveys. A method of the regional division combined with models of two-stage cluster sample is used. In determining the number and composition of areas and farms are respected. The requirements to fall areas and farms in the sample, which reproduce in adequate level structure of agricultural holdings in their specialization are taken into account.

There are an estimated two dispersions, one, called the inter-cluster which represents the magnitude of the dispersion in each planning region, and the second relates to the dispersion inside the districts. It is done in order to achieve the necessary consistency between the farms structure in general aggregation and their structure in the sample. Additionally two indicators to verify the representativeness of the sample are used. The first reflects the degree of corre-

spondence between the average number of farms in one nest of the sample and the average number of farms in one area in the general aggregation. The second indicator measures the degree of correspondence between the variation in chosen nest of the sample and the total variation in the general aggregation. Considering the obtained values of inter-cluster and intra-cluster dispersion, the following rule have used. First, from each planning region is selected this district that most closely approximates the average of farms structure in the region. Additionally selected those districts that have the greatest deviation from the average structure in the relevant planning region. Thus it is ensure that the sample will fall districts sufficient to cover production diversity in the region and will reproduce the structure of farms.

The scope of the sample included farms from 20 districts in the country. Districts whit missing farms in the sample are those whose specialization structure is very similar to the structure of the neighboring districts. The maximum permissible error at guarantee multiplier ( $z = 1.96$ ) is equal to 4.74%. Therefore, it can be argued with guarantee 95% probability that the actual values of the evaluated parameters from the questionnaire information will move within a confidence interval of  $\pm 4.27\%$ .

**Key words:** sample, representativeness, farms, districts, maximum error