

---

# Подход за определяне представителността на извадка при анкетно проучване

Гл. ас. д-р Ангел Саров

*Институт по аграрна икономика – София*

E-mail: angelsarov@abv.bg

## Резюме

Много често в научните изследвания като методически подход за анализ на определени критерии се прилага анкетният метод. След като вече е направен анализ на анкетата и дадена оценка на съответните показатели, възниква въпросът за представителността на извадката и дали получената средна стойност съвпада със средната оценка на генералната съвкупност. В статистиката има разработени методи, които позволяват с помощта на правилата за вероятност да се направят изводи за цялата съвкупност.

В настоящата разработка се апробира подход за определяне представителността на извадка при анкетно проучване. Това се постига чрез екстраполация на извадката към по-общо заключение за цялата популация.

Предложената методика за определяне представителността на извадка при анкетното проучване показва своята приложимост. Подходът позволява възпроизвеждане представителността на резултатите от генералната съвкупност при установяване на управленческата устойчивост на земеделските стопанства.

**Ключови думи:** методика; представителност на извадка; анкетно проучване; оценка на управленческа устойчивост

## Approach for determining the representativeness of a sample in a study

Chief Assist. Prof. Angel Sarov

*Institute of Agricultural Economics – Sofia*

E-mail: angelsarov@abv.bg

**Citation:** Sarov, A. (2021). Approach for determining the representativeness of a sample in a study. *Ikonomika i upravljenie na selskoto stopanstvo*, 66(3), 38-46 (Bg).

## Abstract

Very often in research, as a methodological approach for the analysis of certain criteria, the research method is applied. After analyzing the survey and evaluating the relevant indicators, the question arises about the representativeness of the sample and whether the obtained average value coincides with the average for the general population. In statistics, methods have been developed that allow with the help of probabilistic rules to draw conclusions for the entire population.

This article tests an approach to determining the representativeness of a sample in a study. This is achieved by extrapolating the sample to a more general conclusion for the whole population.

The proposed methodology for determining the representativeness of a sample in the survey showed its applicability. The approach allows the reproduction of the representativeness of the results of the general population in establishing the managerial sustainability of agricultural holdings.

**Keywords:** methodology; sample representativeness survey; governance sustainability assessment

## Увод

Много често в научните изследвания при анализ на определени критерии като методически подход се прилага анкетният метод. След като вече е направен анализ на анкетата и дадена оценка на съответните показатели, възниква въпросът за представителността на извадката и дали получената средна стойност съвпада със средната стойност на генералната съвкупност. Например в монографията „Оценка на управленческата устойчивост за земеделските стопанства в България“ (Саров, А., 2019), която на практика доразвива анализа на предложения „четвърти стълб“ за оценка на устойчивостта по проект на ИАИ „Устойчивост на селското стопанство в България“ (2017–2018), с ръководител проф. Х. Башев, при оценка на управленческата устойчивост на земеделските стопанства възникна въпросът за представителността на направените заключения. Допълнително по темата е достъпна публикация „Оценка на управленческата устойчивост на земеделските стопанства, разположени в планински райони с природни ограничения“ в сп. Икономика и управление на селското стопанство, 64(2), 51-59 (Саров, А., 2019).

За да се определи равнището на устойчивостта на земеделските стопанства в Проекта<sup>1</sup> бе разработена анкетна карта, в която бяха включени 22 въпроса, отнасящи се до управленческия стълб. Анкетното проучване бе проведено през 2018 г. с любезната помощ от страна на НССЗ. Анкетирани бяха мениджъри от 106 земеделски стопанства с различен юридически статут (физически лица; еднолични стопанства; кооперации; търговски дружества и сдружения) и местоположение, покриващо шестте райони на страната (Северозападен; Северен централен; Североизточен; Югозападен; Южен централен и Югоизточен). Анкетата обхваща критерии по различна специализация на стопанствата – специализирани в производството на полски

култури; зеленчуци, цветя и гъби; трайни насаждения; тревопасни животни; свине, птици и зайци; смесено растениевъдни; смесено животновъдни; смесено растениевъдно-животновъдни; пчелари, както и различен размер на стопанството, съответно различен размер на използваната земеделска земя.

В случая са използвани ограничени количества данни, които резонно поставят въпроса дали направените заключения са релевантни за цялата съвкупност.

## Подход за определяне представителността на извадка при анкетно проучване и получаване на средни стойности в различни групировки

В настоящата разработка се апробира подход за определяне представителността на извадка при анкетно проучване (Иванов, Б., стр. 12-20 в Башев, Хр. и кол., 2021)<sup>2</sup>. Известно е, че в статистиката има разработени методи, които позволяват с помощта на правилата за вероятност да се направят изводи за цялата съвкупност. Това се постига чрез екстраполация на извадката към по-общо заключение за цялата популация.

Проблемът е, че статистическите изводи могат да се прилагат само за популацията, от която са получени резултатите на извадката. Статистическите методи позволяват да се направят изводи за това, което вероятно би се случило, ако се повторят анкетите още много пъти в генералната съвкупност.

Теоретично погледнато, за да се подготви представителна за страната извадка от земеделски производители, които да бъдат анкетирани, получените резултати следва да се екстраполират за генералната съвкупност с избрана степен на гаранционна вероятност. Тъй като вече анкетата е осъществена

<sup>1</sup> ИАИ, Проект „Устойчивост на селското стопанство в България“ (2017-2018), с ръководител проф. д-р Х. Башев, ИАИ, ССА.

<sup>2</sup> Храбрин Башев, Божидар Иванов, Диляна Митова, Иван Боевски, Петър Маринов, Ангел Саров, Даниела Цвяткова, Красимир Костенаров, Димитър Ванев, (2021). „Механизми и форми на управление на агро-еко-системните услуги в България“. Издателство: Институт по аграрна икономика, Печатница „Авангард Прима“, ISBN 978-954-8612-32-6.

на със 106 земеделски производители, съответно размерът на извадката е равен на 106, то в конкретния случай за екстраполация ще се подходи обратно. До момента на база приетата методика е изчислено, че средната стойност на извадката е с коефициент 0,41. За да се подходи „чисто“ научно и, за да бъдат елиминирани евентуални съмнения относно интерпретацията на получените стойности, ще бъде изчислена стандартната грешка на извадката, след което ще стане известен новият ѝ размер.

За да се определи средната стохастична грешка на оценката, се изчислява т.н. максимално допустима грешка, при което се взема предвид доверителната вероятност и съответният гаранционен множител  $Z$ . Известно е, че максималната грешка е правопрпорционална на стандартната грешка и на доверителния коефициент. Все пак трябва да се има предвид, че ако се приеме по-голяма вероятност, то ще се увеличи максималната грешка и съответно диапазонът на доверителния интервал се увеличава. Всичко това прави изключително противоречиво, от една страна, максималната грешка да е колкото е възможно по-малка и заключението да бъде с по-голяма вероятност, от друга. В случая се работи с анкета, в която са попаднали едва 106 земеделски производители на фона на генералната съвкупност в страната. При една подобна ситуация, при която извадката ( $n$ ) е малка, спрямо генералната съвкупност ( $N$ ), то стандартната грешка би клоняла към 1.

Първата стъпка е да се определи доверителният интервал. Той показва, каква част от генералната съвкупност се наблюдава и каква е представителността, което е показателно за представителността на резултатите.

Избира се доверителен интервал 95%. На практика това означава, че с 95% вероятност доверителният интервал ще съдържа истинската средна стойност на популацията.

Използваните стойности са 95% гаранционна вероятност на получените оценъчни резултати, където  $Z = 1,96$ .

Приема се нормално разпределение на съвкупността, което означава, че крайните слу-

чай са само 2,5% в единия край и толкова в другия. Преобладаващата част от случаите са около средната стойност на групировката.

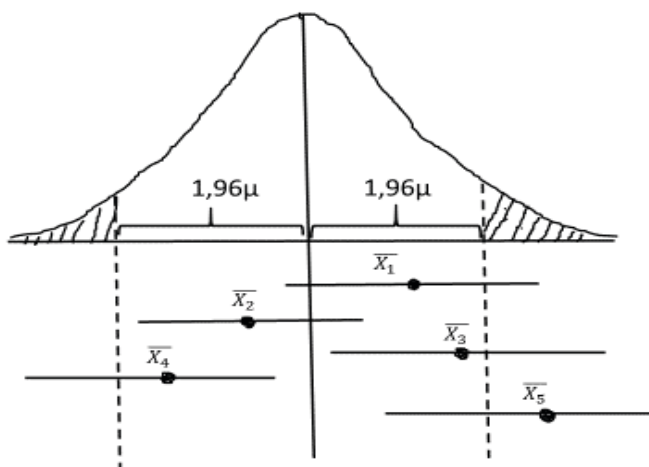
Със сигурност се знае, че средната стойност на цялата популация има една стойност. Приема се, че в конкретния случай не се знае каква е тя. За разлика от това, доверителният интервал, който се изчислява, зависи от данните, които са събрани. Ако се повтори анкетата с други земеделски стопанства, новият доверителен интервал почти със сигурност ще бъде различен.

Ако се проведе тази анкета няколко пъти, интервалите на доверие не биха били еднакви, но може да се очаква, че 95% от тях ще съдържат средната стойност. Може да се очаква, че 5% от доверителните интервали няма да включват средната стойност за популацията. За да се интерпретира доверителният интервал на средната стойност, следва да се приеме, че всички стойности са взети независимо и произволно извадени от популация, чиито стойности са групирани според разпределението на Гаус. При приемане на тези предположения има 95% вероятност доверителният интервал да съдържа действителната средна стойност.

В настоящата анкета, както стана ясно, са интервюирани 106 земеделски стопанства. Доверителните интервали имат съответно разпределение. Ако е известна средната стойност на генералната съвкупност  $\bar{X}_0$ , то тогава тя се намира в едни интервали, а в други липсва. Ако са направени например 5 отделни анкети от тази популация с различни интервюирани земеделски стопанства, то интервалите на  $\bar{X}_1; \bar{X}_2; \bar{X}_3; \bar{X}_4$  съдържат  $\bar{X}_0$ , но един интервал, а именно  $\bar{X}_5$  – не го включва. Или, в случая, от 5 интервала 95% ще попаднат в  $\bar{X}_0$  (фиг. 1).

По подобен начин може да се тълкува доверителният интервал на извадката. Т.е. 95% е вероятността доверителният интервал на средната стойност на извадката да съвпада със средните стойности на анкетираните 106 стопанства.

На практика доверителните интервали могат да бъдат изчислени с различна степен на



Фиг. 1. Доверителен интервал – разпределение на средните в общата съвкупност

Fig. 1. Confidence interval – distribution of the average in the population

Източник: Собствена фигура. / Source: Own figure.

доверие. В случая е избран 95%. Следва да се има предвид, че 99% доверителен интервал е по-широк от 95%, а съответно 90% доверителен интервал е по-тесен, което е логично. В случай, ако се търси повече гаранция, че доверителният интервал съдържа действителната стойност, тогава интервалите следва да бъдат по-широки. Ако се търси 100% гаранция, че интервалът съдържа цялата популация, той трябва да съдържа всяка възможна стойност, което означава, че ще бъде много широк. Ако се търси например 60% сигурност, че интервалът съдържа истинската стойност, тогава той ще бъде много по-тесен.

Доверителният интервал на средната стойност показва каква е точността на определената средна стойност на извадката.

Например изчислената средна стойност на извадката от 106 стопанства ( $n = 106$ ) е малко вероятно да се равнява на средната за популацията. В случая се приема, че извадката е малка и променлива, а средната стойност на извадката най-вероятно ще бъде различна от средната на цялата популация. Ако извадката е голяма, например ( $n = 10\,600$ ) и има малко разсейване, то средната стойност на извадката вероятно ще бъде много близка до средната стойност на популацията. Известно е, че

статистическите изчисления съчетават размера на извадката и вариабилността (стандартно отклонение), за да генерират доверителен интервал за средната стойност на популацията.

За да се определи размерът на извадката, се ползва следната формула на Cochran (1963):

$$SS = \frac{Z^2 * p * (1 - p)}{c^2} \quad (1)$$

Където:

SS – размерът на извадката;

Z – Z value (в случая 1,96 за 95% доверителен интервал);

p – избран процент на избор, изразен като десетичен знак (0,5 – използва се за необходимия размер на извадката);

c – средна грешка.

При извадка, когато тя се обособи в две групи, в които се вижда, че всички белези на анкетираните са в определено съотношение, това съотношение на случаите в извадката се има предвид при процент на направените избори. Обикновено се приема, че този коефициент е 0,5.

В случая, от решението на формулата ще се изчисли средната допустима грешка.

След това се определя новият размер на извадката по формулата на Israel (2003):

$$\text{new } SS = \frac{SS}{1 + \frac{SS-1}{N \text{ pop}}} \quad (2)$$

Където:

new SS – новият размер на извадката;

SS – размерът на извадката;

N pop – популация.

При определяне размера на анкетираните стопанства се ползва икономическият им размер (ИР). В анкетата попадат предимно физически лица, чийто земеделски стопанства са с икономически размер на 1 ЗС (СПО – хил. евро) до 4000 евро, т. е. попадат в групата на малки и средни стопанства. Това са полупазарни стопанства, които на практика са сред най-уязвимата група български производители. Именно те преобладават в целевата група на интервюираните. Според данни от МЗХГ, отдел „Агростатистика“, за периода 2010–2016 г. малките и средните стопанства в България с размер до 4000 евро намаляват чувствително, като от 314 578 през 2010 г., достигат 139 854 през 2016 г. В същото време стопанствата с размер над 8000 евро за този период се увеличават с близо 9 000 броя, съответно от 29 358 през 2010 г., до 38 205 през 2016 г.

Според Доклад „Анализ на състоянието на селското стопанство и хранително-вкусовата промишленост, SWOT анализ“ – ИАИ, (2020), се определя, че подкрепените стопанства чрез директните плащания са предимно с икономически размер над 4 000 евро. Това още повече засилва необходимостта от настоящето изследване при вземането на решения от полисимейкърите за установяване на проблемите пред развитието на тази група стопанства.

В случая ще се приеме, че броят на земеделските стопанства с ИР до 4000 евро (популацията – *pop*), ще е 150 000.

След заместване във формулите се получава, че стандартната грешка е 10% (0,1), което означава, че при доверителен интервал от 95% получените резултати (средната оценка на извадката) следва да се интерпретират с  $\pm 10\%$ .

Според изчисленията, новият размер на извадката е в границите на 105,9259 или приблизително 106.

По-нататък изследването продължава с проверка на хипотези, за да се проследи хомогенността на извадката. Колкото по-разсеяни са показателите, толкова по-ненадеждни са резултатите от нея. Проверката на хипотези представлява подход, при който формираната извадка се разделя по определен признак, в случая географски. Целта е да се видят разликите в получаваните средни стойности и отклонения в двете подизвадки. За целта се определя един релевантно корекционен коефициент, който да послужи за аджестиране на оценките на устойчивостта на база първичните резултати.

Следващите стъпки са намиране на новите средни стойности и стандартна грешка на две подгрупи. При определянето на районите са изпълнени критериите, гарантиращи в различните райони да попаднат стопанства, в които се възпроизвежда в достатъчно висока степен структурата на земеделските стопанства по тяхната специализация и възраст на управителя. В първата група се избират 3 от 6-те района на планиране при формирането на извадката в страната, а именно: Северо-западен, Северен централен и Североизточен район. В направената анкета, те са представени от 56 земеделски стопанства. Във втората група попадат следващите 3 района: Югозападен, Южен централен и Югоизточен район, в която са включени 50 стопанства.

Под внимание се взема обстоятелството, че изводите на подобно сравнение се отнасят до цялата съвкупност, въпреки че данните, с които се работи, обхващат само извадката. Формулират се предположения – хипотези, като се прави проверка дали данните от извадката се потвърждават или отхвърлят.

Хомогенността на извадката ще се определи на база наличие на разлики в получените средни стойности и отклонения в двете подизвадки. Конструират се две предположения – „не, двете подгрупи не са хомогенни“ и „да, двете подгрупи са хомогенни“. Тези две възможности се изразяват чрез формулиране на две статистически хипотези:

- Нулева хипотеза ( $H_0$ ) – тя предполага, че няма статистически значима разлика в срав-

няваните статистически показатели (средните стойности и стандартните отклонения в двете подизвадки), спрямо тези на генералната съвкупност.

- *Алтернативна хипотеза ( $H_1$ )*, според която, има статистически значима разлика в сравняваните статистически показатели (средните стойности и стандартните отклонения в двете подизвадки), спрямо тези на генералната съвкупност.

Проверката на статистически хипотези има вероятностен характер, поради факта че изследванията са репрезентативни, т.е. изводите за съвкупността се правят въз основа на изследване на малка извадка. Приемането или отхвърлянето на нулевата хипотеза се прави с определена вероятност, като в същото време се допуска възможността за грешка.

За сравнение на двете подгрупи избираме да приложим дисперсионен анализ или т.нар. ANOVA test., който се отнася към методите за изследване на връзки и зависимости. С негова помощ ще се установи дали между двете подгрупи в извадката съществува хомогенност и какво е разпределението вътре в подгрупите. Най-общо казано, ще се тестват двете групи, за да се провери дали има сходства между тях и дали установените стандартна грешка и отклонение са значими. Чрез проверка на хипотези може да се прецени доколко влиянието на даден фактор-причина или на група фактори-причини е статистически значимо или не. ANOVA ще помогне да се установи дали да се отхвърли нулевата хипотеза или приеме алтернативната хипотеза.

Приети са следните стойности за гаранционната вероятност (P) и равнище на значимост (a):

- при P = 95%, съответства a = 0,05 (5% възможна грешка).
- при P = 99%, съответства a = 0,01 (1% възможна грешка).
- при P = 99,9%, съответства a = 0,001 (0,1% възможна грешка).

При положение, че P < 95%, a > 0,05, за вярна следва да бъде приета нулевата хипотеза. В обратния случай – алтернативната.

Накратко стъпките за проверка на статистическите хипотези са следните:

На първо място се формулират нулевата и алтернативната хипотеза. Според нулевата ( $H_0$ ) хипотеза няма разлика в сравняваните показатели. Ако емпиричните данни показват различия, то те се дължат на случайни явления. Противоположно е твърдението на алтернативната хипотеза, според която наблюдаваните в емпиричните данни различия са резултат на закономерности.

Вторият етап е да изберем подходящ критерий за проверка на хипотезата.

За целта е необходимо да се познават условията, на които трябва да отговарят променливите.

Третата стъпка е изчисляване на емпиричната стойност на критерия ( $K_{emp}$ ).

Емпиричната стойност на критерия се изчислява ( $t_{emp}$ ) по формула:

$$t_{emp} = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\frac{S_1^2(n_1 - 1) + S_2^2(n_2 - 1)}{n_1 + n_2 - 2} \cdot \frac{n_1 + n_2}{n_1 n_2}}} \quad (3)$$

Където:

$\bar{X}_1$  и  $\bar{X}_2$  – средни стойности на извадките;  
 $S_1^2$  и  $S_2^2$  – стандартни отклонения на извадките;

$n_1$  и  $n_2$  – обеми на извадките.

Определя се таблична стойност на критерия ( $t_a$ ), според таблично приложение, в зависимост от равнището на значимост (a) и степените на свобода  $k = n_1 + n_2 - 2$ .

Сравняваме емпиричния с теоретичния критерий. Ако  $t_{emp} > t_a$ , отхвърляме  $H_0$ .

Възможностите са две:

- **Ако емпиричната стойност е по-малка от табличната**, за вярна се приема нулевата хипотеза ( $H_0$ ).

- **Ако емпиричната стойност е по-голяма или равна на табличната** – нулевата хипотеза ( $H_0$ ) се отхвърля и за вярна се приема алтернативната хипотеза ( $H_1$ ).

В такъв случай:

- Ако гаранционната вероятност е по-малка от възприетата (95% или 99%) – за вярна се приема нулевата хипотеза ( $H_0$ ).

• Ако гаранционната вероятност е равна или по-голяма от 95% (99%) – нулевата хипотеза се отхвърля и за вярна се приема алтернативната хипотеза.

Петата стъпка е вземане на решение. За целта табличната (критичната) стойност на критерия се сравнява с емпиричната (изчислена по данни от извадката). Именно това сравнение дава основание да се направи извод коя от хипотезите (нулевата или алтернативната) е вярна.

В нашия случай приемаме за критерий ( $t_a$ ) 3,2905, при степен на свобода  $Z$ , с 99,9% точност, на която съответства  $\alpha = 0,001$  (0,1% възможност за грешка):

➤ Когато  $t_{emp}$  е по-малък от 3,29, се приема, че няма статистически значима разлика в сравняваните статистически показатели (средните стойности и стандартните отклонения в двете подизвадки), спрямо тези на генералната съвкупност.

➤ Когато  $t_{emp}$  е по-голям от 3,29, тогава се приема, че разликите са значими.

След това се изчисляват новите средни стойности ( $\bar{X}_1$  и  $\bar{X}_2$ ), съответно за двете групи. След направените изчисления, средната стойност на 1-ва група е 0,417, а на втората – 0,387. Прави впечатление, че средната стойност на втората група е под средната  $\bar{X}_0$  на общата извадка, която е 0,4105. Това означава, че най-вероятно значително по-голямата част от земеделските производители имат по-ниска средна от средната стойност на генералната съвкупност. За да се постигне необходимото съответствие между оценката на управленческата устойчивост на земеделските стопанства от генералната съвкупност и тази в извадката, са изчислени двете дисперсии – съответно в двете групи стопанства.

За изчисление на стандартната грешка на двете групи се използва формулата:

$$SE = \frac{\sigma_{1*(1-R)}}{\bar{X}_{total}} - \frac{\sigma_{2*(1-R)}}{\bar{X}_{total}}, \quad (4)$$

Където:

SE – стандартна грешка;

$\sigma_{1;2}$  – стандартното отклонение на двете групи;

$R$  – представителността на извадката;

$\bar{X}_{total}$  – средната стойност на извадката.

На първата група стопанства вероятността за стандартната грешка е 0,190931, а във втората – 0,147898. След изчисление на SE, която е равна на разликата между двете стойности, съответно 0,043033, което на практика е показател за хомогенност или различия между двете групи.

При така направените изчисления на извадката, стандартната грешка е равна на 4,3%. Съответно максималната грешка ( $\Delta_0$ ) при гаранционна вероятност ( $z = 1,96$ ) ще е равна на  $\sqrt{4,3 \times 1,96} = 4,07\%$ . В този случай с 95% гаранционна вероятност може да се твърди, че действителните стойности на параметрите от анкетната извадка ще се движат в границите на доверителен интервал  $\pm 4,07\%$ .

Според изчисленията  $F$  ANOVA е с параметри  $3,67 > 3,29$ . Това показва, че има статистически значима разлика в сравняваните статистически показатели (средните стойности и стандартните отклонения в двете подизвадки), спрямо тези на генералната съвкупност. Следователно се отхвърля нулевата хипотеза  $H_0$ , която отрича наличие на статистически значима разлика в сравняваните статистически показатели (средните стойности и стандартните отклонения в двете подизвадки, и се приема алтернативната хипотеза  $H_1$ .

Следваща стъпка е намиране на стандартното отклонение на извадката, което се изчислява чрез формулата, а именно:

$$SD_{total} = \frac{SD_{\bar{X}_1} + SD_{\bar{X}_2}}{2} \quad (5)$$

Преди това обаче ще е необходимо да се изчислят съответно стандартните отклонения на двете групи. Това се получава с помощта на следните две формули:

$$SD_{\bar{X}_1} = \frac{\bar{X}_1 * SE}{\bar{X}_{total}} \quad (6)$$

$$SD_{\bar{X}_2} = \frac{\bar{X}_2 * SE}{\bar{X}_{total}} \quad (7)$$

Стандартното отклонение на извадката е 0,042135. То е необходимо, за да се изчисли новата средна стойност на извадката.

За да са коректни изчисленията, следва да се намери медианата на извадката. Между 53 и 54 номер на извадката средните стойности са в границите между 0,336124 и 0,339352.

Тук има две възможности: Първо – Ако средната на извадката е по-висока от медианата, за да се изчисли новата средна, по-коректно е да се ползва първата формула. Второ – Ако средната на извадката е по-ниска от медианата, препоръчителна е втората. В конкретния случай, тъй като медианата е доста по-ниска от средната на извадката, то се прилага втората формула.

$$\text{New } \bar{X} \text{ IF } \bar{X}_{total} > \text{Median than} \\ \bar{X}_{total} = [(\bar{X}_1 - \bar{X}_1 * SD_{total}) + (\bar{X}_1 + \bar{X}_2 * SD_{total})] / 2 \quad (8)$$

$$\text{IF } \bar{X}_{total} < \text{Median than} \\ \bar{X}_{total} = [(\bar{X}_2 - \bar{X}_2 * SD_{total}) + (\bar{X}_2 + \bar{X}_1 * SD_{total})] / 2 \quad (9)$$

Where  $\bar{X}_1 > \bar{X}_2$

Получените коефициенти в резултат на използваните допълнителни два критерия за представителност на извадката са следните:

Новата средна стойност на извадката е 0,3874, което спрямо резултатите от първоначалните изчисления (0,4105) се отклонява с 0,0231. Имайки предвид стандартната грешка, новата средна стойност на извадката попада в границите  $\pm 10\%$ .

За да се изчисли доверителният интервал на получената нова средна, ще се използва следната формула:

$$\bar{x} \pm z \frac{s}{\sqrt{n}} \quad (10)$$

Където:

$\bar{x}$  – новата средна на извадката;

Z – доверителен интервал 95%, Z = 1,96;

s – стандартната грешка;

n – размер на извадката.

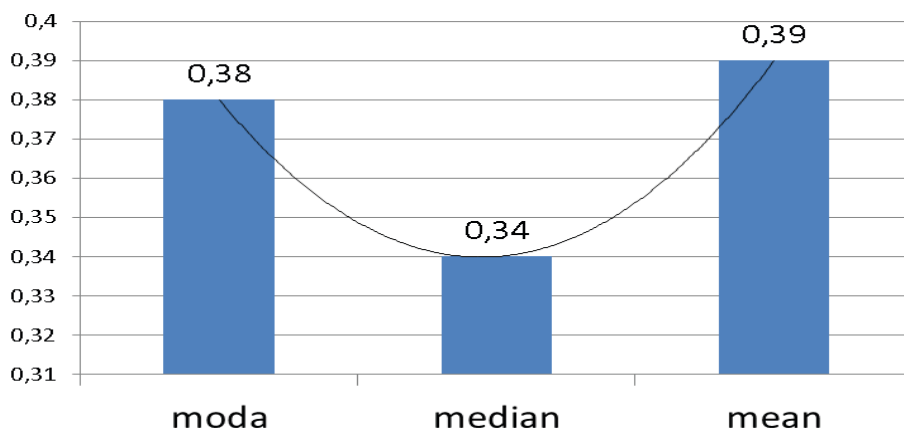
При доверителен интервал 95%, то съответно средната стойност на извадката попада в диапазона долна и горна граница – между 0,371 и 0,391.

Както се вижда, двата показателя попадат под границата на 0,4, според която, земеделските стопанства в България получават незадоволителна оценка на управленческата устойчивост.

Информация за формата на разпределението е представена графично в мерките на основната тенденция (*средна, медиана, мода*) (фиг. 2). Те трябва да се приемат като вид средни стойности и отразяват тенденцията в разположението на центъра на честотното разпределение.

На фиг. 2 се вижда, че Median < moda < mean, което означава, че разпределението е несиметрично спрямо средната стойност.

Модата  $M_0$  се определя като показателят с най-голяма честота в разпределението. При



Фиг. 2. Разпределение на Median, Moda, Mean

Фиг. 2. Distribution of Median, Mode, Mean

Източник: Собствени изчисления. / Source: Own calculations.



групирани данни за *мода* се приема средата на класа с най-голяма честота. В случая има многомодално разпределение. Това е така, защото има няколко стойности, които се срещат еднакво често, с честота по-голяма от тази на останалите. Стойността 0,38 на *модата* е изчислена като средното аритметично на тези стойности. Ниската стойност на *медианата* означава, че управленческата устойчивост на една значителна част от земеделските стопанства в България е поставена пред изпитание. Едновременно с това прави впечатление, че коефициентът 0,38 на *модата* се доближава до *средната* на извадката (0,39), попадайки в обхвата на получената стандартна грешка.

## Заклучение

Предложената методика за определяне представителността на извадка при анкетно проучване покрива критериите на поставената задача. Методиката възпроизвежда в най-голяма степен представителността на резултатите от генералната съвкупност при установяване на управленческата устойчивост на земеделските стопанства. В нейна подкрепа е и доказателството, че не може да се предположи доколко избраните и попаднали в извадката стопанства отговарят на действителното разпределение на стопанствата от общата популация.

Новата средна на извадката е 0,39, което спрямо резултатите от първоначалните из-

числения (0,41) се отклонява с 0,02. Предвид стандартната грешка, новата средна на извадката попада в границите  $\pm 10\%$ . Този коефициент е приемлив от гледна точка на факта, че той не би могъл да промени в голяма степен получения резултат и направената оценка на управленческата устойчивост на земеделските стопанства в България.

## Литература

**Саров, А.** (2019). Оценка на управленческата устойчивост за земеделските стопанства в България. Издателство „Авангард Прима“.

**Башев, Х., Иванов, Б., Митова, Д., Боевски, И., Маринов, П., Саров, А., ... & Ванев, Д.** (2021). Механизми и форми на управление на агроecosистемните услуги в България. Издателство: Институт по аграрна икономика, Печатница „Авангард Прима“, ISBN 978-954-8612-32-6.

**Cochran, W.G.** (1963). *Sampling Technique*. 2nd Edition, John Wiley and Sons Inc., New York.

**Sarov, A.** (2019). Assessment of the governance sustainability of farms located in mountain areas with natural constraints. *Bulgarian Journal of Agricultural Economics and Management*, 64(2), 51-59.

ИАИ. (2020). Анализ на състоянието на селското стопанство и хранително-вкусовата промишленост, SWOT анализ.

ИАИ. (2019). Проект „Устойчивост на селското стопанство в България“ (2017–2018), с ръководител проф. д-р Храбрин Башев.

МЗХГ. (2016). Отдел „Агростатистика“. Преброяване на земеделските стопанства (2010–2016).