Гл. ас. д-р Поля Ангелова

МОДЕЛИРАНЕ РАЗПРЕДЕЛЕНИЕТО НА ЗЕМЕДЕЛСКИТЕ СТОПАНСТВА В БЪЛГАРИЯ ЗА ПЕРИОДА 1897 – 2005 Г.*

Въз основа на информация от преброяванията на поземлената собственост и земеделските стопанства в България за периода 1897 – 2005 г. са предложени възможни варианти за моделиране на емпиричното разпределение на земеделските стопанства според размера на стопанисваната земя. Използвани са три функции при моделирането, чийто графичен образ се доближава до емпиричните криви на разпределението: експоненциалното разпределение, разпределението на Парето и ляво отсеченото нормално разпределение. Направен е преглед на основните характеристики на посочените разпределения. Оценени са параметрите на трите функции чрез метода на моментите, като е аргументиран неговият избор. Установени са теоретичните честоти за отделните наблюдения и е оценено съответствието между емпиричните и теоретичните разпределения. Дадени са изводи относно особеностите на разпределението, свързани с крайно асиметричната му форма, и влиянието им върху състоянието на съвременното българско земеделие.

JEL: C16; C51; Q15

Присъединяването на България към Европейския съюз и възстановяването на традиционните й позиции на селскостопански износител в международния стокообмен поставя съществения въпрос за адекватността на разпределението на земеделските стопанства според размера на земята към изискванията на Общата селскостопанска политика на Общността. Разпределението според размера на стопанисваната земя се формира под влиянието на множество условия – икономически, социални, демографски и т.н., и обединява факторите на производството в селското стопанство – земя, аграрен капитал и земеделски труд. То се оказва определящо по отношение на: производствените възможности на земеделските стопанства; равнището на използване на селскостопанската техника; разпределението и използването на работната сила; професионалния и социалния състав на заетите в селското стопанство; аграрната политика и ролята на държавата за подпомагане на земеделските производители и т.н.

Функцията на разпределението, известна в теорията на статистиката като функция на плътността на разпределението, е една от основните статистически характеристики, които свързват по определен начин единиците на съвкупността с отделните варианти на признака. Нейното моделиране е етап от статистическия анализ на динамиката на вариационните разпределения.[1]

---

* Статията е част от извършено от автора статистико-икономическо изследване на динамиката на вариационните разпределения на земеделските стопанства в България за периода 1897 – 2005 г. Информационното осигуряване на изследването се базира на официалната статистическа информация от преброяванията на поземлената собственост и земеделските стопанства за същия период.

[1] *Величкова, Н.* Статистически методи за изучаване и прогнозиране на развитието на социално-икономическите явления. С.: "Наука и изкуство", 1981, с. 370 - 400.

Избор на функция за моделиране на разпределението

Историческият преглед на състоянието и особеностите на статистическите наблюдения за изследвания период ни помогна да обосновем необходимите изисквания - за еднородност, съпоставимост и сравнимост на данните. Изследването на състоянието и структурата на съвкупността, от една страна, и динамиката на обобщаващите характеристики, от друга, осигури необходимата основа за избор на подходяща функция за моделиране и оценка на параметрите.

Данните за разпределението на земеделските стопанства според размера на стопанисваната земя са представени във вариационно-динамична таблица (табл. 1).

Таблица 1

Динамика на разпределението на земеделските стопанства според размера на стопанисваната земя в България за периода 1897 – 2005 г. (бр.)

Групи стопанства според размера на земята (дка)	1897 г.	1908 г.	1926 г.	1934 г.	1946 г.	2001 г.	2003 г.	2005 г.
До 10	257 335	293 750	89 040	119 627	154 563	1 210 431	506 305	385 647
Над 10 “ 20	106 357	131 148	92 895	119 790	154 728	136 204	86 666	73 940
“ 20 “ 30	75 089	86 500	89 837	116 967	149814	49 815	24 412	22 593
“ 30 “ 40	60 056	68 346	82 592	107 817	137 586	24 273	10 776	9787
“ 40 “ 50	50 220	57 772	73 155	94 904	115 729	12 525	5701	6122
“ 50 “ 60	37 010	42 665	60 753	72 894	86 829	8099	3725	4181
“ 60 “ 70	37 010	42 665	50 222	56 732	64 702	4456	2318	1978
“ 70 “ 80	29 803	34 822	40 301	43 288	45 849	2558	1547	1667
“ 80 “ 90	22 596	26 979	33 229	33 621	33 503	1852	1125	1086
“ 90 “ 100	22 596	26 979	25 936	25 346	23 977	1037	901	1099
“ 100 “ 150	55 509	67 610	70 605	62 488	51 618	2876	2657	3162
“ 150 “ 200	22 097	26 718	24 000	18 745	12 936	793	1217	1306
“ 200 “ 300	14 913	17 304	13 346	9623	5740	620	1256	1463
“ 300 “ 400	4338	5017	3025	1927	978	306	697	751
“ 400 “ 500	1770	1933	857	539	250	139	469	586
Над 500	2941	3159	820	561	270	622	5036	5165
Всичко	799 642	933 367	750 613	884 869	1 039 072	1 456 606	654 808	520 533

Източник: Статистика на поземлената собственост през 1908 г. С., Главна дирекция на статистиката, 1914; Преброяване на земеделските стопанства в Царство България на 31 декември 1926 г. С., Главна дирекция на статистиката, 1935; Преброяване на земеделските стопанства 1934 г. С., Главна дирекция на статистиката, 1942; Преброяване на земеделските стопанства през м. август 1946 г. Месечни известия на Главна дирекция на статистиката, 1/1947; Преброяване на населението, жилищния фонда и земеделските стопанства през 2001. Т. 5 Земеделски стопанства. С., НСИ, 2003; Преброяване на земеделските стопанства’2003 и Структурна анкета за стопанската 2004/2005 г. С., МЗГ, Агростатистика, 2005; изчисления на автора.

Моделирането на изследваното емпирично разпределение се извършва чрез използване на различни теоретични разпределения. В конкретното изследване емпиричното разпределение на земеделските стопанства според размера на земята има крайно асиметрична форма и изборът на подходяща функция за неговото моделиране предизвика редица трудности и проблеми при апроксимирането му чрез използване на някои от известните теоретични разпределения.

Анализът на обобщаващите характеристики на разпределението и графичният образ на кривите показват, че през изследвания период са настъпили количествени изменения в разпределението на земеделските стопанства според размера на земята. Тези промени предизвикват и качествени изменения във функцията, която съответства на дадено разпределение. Постепенното намаление на честотите при нарастване на случайната променлива е характерно за първите две наблюдения. През следващите три наблюдения се забелязва, макар и слабо изразена, върхна източеност на кривата на разпределението, като модата се локализира във втората или третата група. В последните три съвкупности кривата отново се връща в първоначалното си положение, но с много рязко намаление на честотите още при ниските значения на признака. На фиг. 1 са представени полигоните на разпределението на земеделските стопанства през изследвания период.

Крайно асиметричната L-образна форма на разпределението най-общо наподобява експоненциална крива. Графичният образ на емпиричните криви на разпределението дава основание да се предпочетат три функции при моделиране на разпределението на земеделските стопанства според размера на земята:

- експоненциално разпределение;
- разпределение на Парето;
- ляво отсечено нормално разпределение.

Експоненциалното разпределение[2] е подходящо за моделиране на явления, които се изменят (нарастват или намаляват) с постоянни темпове. В случая с нарастването на значенията на признака размер на земята броят или относителният дял на стопанствата намалява. Теорията на вероятностите и статистиката отнасят експоненциалното разпределение към групата на непрекъснатите разпределения. Функцията на плътността на разпределението намира израз в следната формула:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, x \geq 0 \\ 0, x \langle 0 \end{cases},$$

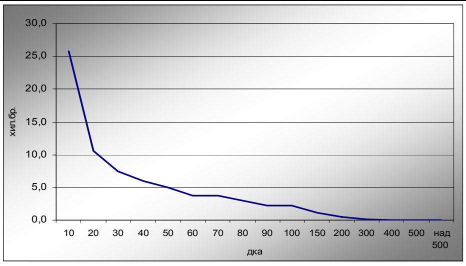
където: x_i е случайна променлива (размерът на земята) и приема винаги положителни стойности;

λ >0 – параметър на разпределението.

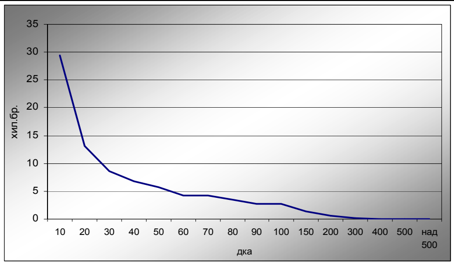
[2] Вж. по-подробно Exponential Distribution. Engineering Statistics Handbook, http://www.itl.nist.gov/div898/handbook/eda/section3/eda3667.htm; *Balakrishnan, N., A. P. Basu*. The Exponential Distribution: Theory, Methods, and Applications. New York: Gordon and Breach, 1996; Экспоненциальное (показательное) распределение. Курс теории вероятностей. http://www.exponenta.ru.

Фигура 1

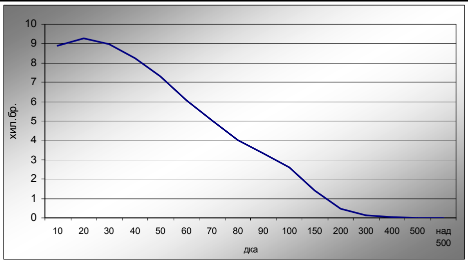
Емпирични криви на разпределението



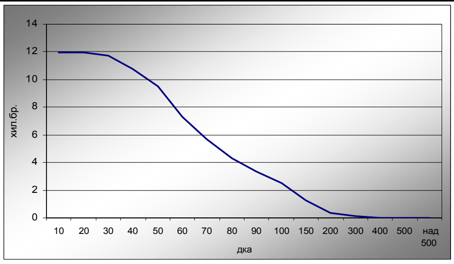
Разпределение на поземлените притежания през 1897 г.



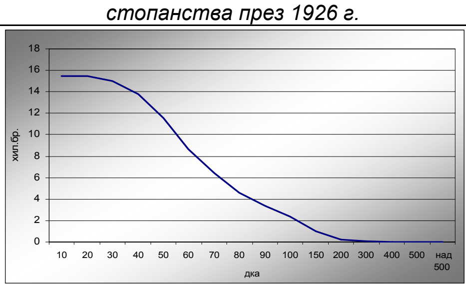
Разпределение на поземлените притежания през 1908 г.



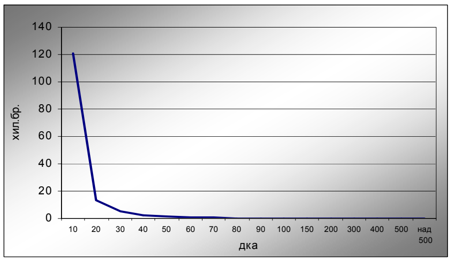
Разпределение на земеделските стопанства през 1926 г.



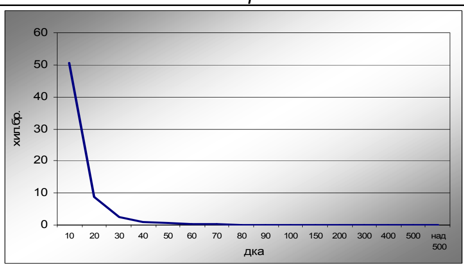
Разпределение на земеделските стопанства през 1934 г.



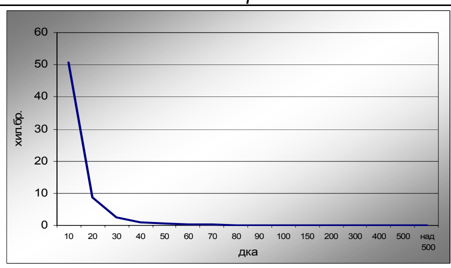
Разпределение на земеделските стопанства през 1946 г.



Разпределение на земеделските стопанства през 2001 г.



Разпределение на земеделските стопанства през 2003 г.



Разпределение на земеделските стопанства през 2005 г.

Разпределението е дефинирано в интервала [0, ∞). Математическото очакване $E(X)$ на експоненциалното разпределение е:

$$E(X)=\int_0^{+\infty} x\lambda\ell^{-\lambda x}dx=\frac{1}{\lambda}.$$

Установяването на теоретичните честоти при моделирането на емпиричното разпределение налага допълнителна параметризация. Функцията на плътността на разпределението може да се представи по следния начин:

$$f(x,\beta)=\begin{cases}\frac{1}{\beta}e^{-\frac{x}{\beta}}, x\geq 0;\\ 0, x\langle 0\end{cases}$$

откъдето става ясно, че $\beta=\frac{1}{\lambda}$ е характеристика за мащаба на разпределението, реципрочна на параметъра λ. Тъй като $E(X)=\frac{1}{\lambda}$, следва, че $E(X)=\beta$.

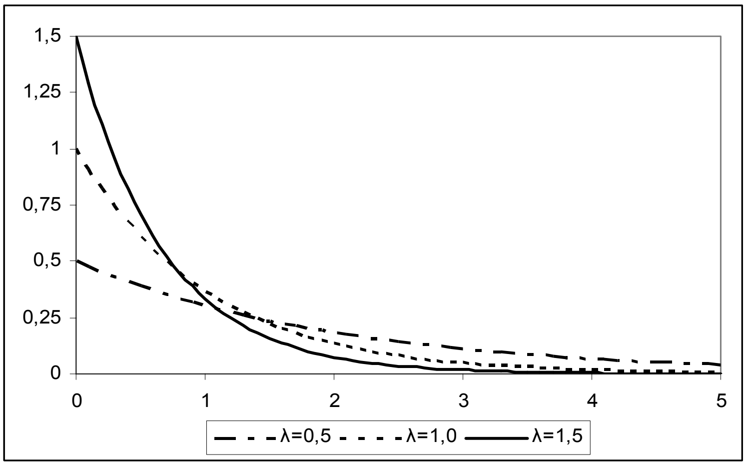
Параметърът β е оценката на средната аритметична величина. В резултат от това се оказва, че интерпретацията на параметъра λ е реципрочната стойност на средната аритметична. Положението на кривата се определя от числовата стойност на λ, който е по-малък от единица ($\frac{1}{\overline{\overline{X}}}$<1) и при високите значения на x_i клони към 0.

Тъй като средноквадратичното отклонение е равно на $\sigma=\sqrt{\frac{1}{\lambda^2}}=\frac{1}{\lambda}=\beta$, със средната аритметична може да се оцени и разсейването на експоненциалното разпределение.

Графично плътността на разпределението е представена на фиг. 2. Както се вижда, кривата е крайно асиметрична и нейното положение се определя от λ. На най-голямата ордината, равна на λ, съответства най-малкото значение на променливата $x=0$. С нарастването на x ординатите постепенно намаляват и при високите значения на $x\to+\infty$ плътността на разпределението клони към нула.

Фигура 2

Функция на плътността на експоненциалното разпределение за различни стойности на параметъра λ



Близка по своя геометричен образ до емпиричната крива на разпределението на земеделските стопанства е *кривата на Парето.*[3] Разпределението на Парето описва статистически закономерности в т. нар. отсечени съвкупности, от които са отделени единици със значения на признака под определен фиксиран размер.[4] Освен близкия графичен образ това обстоятелство е също аргумент за избора на кривата на Парето при моделиране на разпределението. При наблюдението на земеделските стопанства винаги има "праг" по отношение на стопанисваната земя (независимо че е различен за различните наблюдения).

Ако X е случайна променлива с разпределение на Парето, то вероятността x_i да е по-голяма от дадено значение се определя от интегралната функция на разпределението:

---

[3] Вилфредо Парето (1848-1923) - италиански икономист, установил закономерност в разпределението на населението по доходи. За повече подробности вж. *Weber, C., E*. More on Slutsky's Equation as Pareto's Solution. - History of Political Economy. Vol. 31, N 3, Fall 1999, p. 575-586; Vilfredo Pareto. Encyclopædia Britannica, http://www.britannica.com/eb/article-9058449 и др.

[4] Вж. *Ланге, О*. Введение в эконометрику. Москва: "Прогрес", 1964, с.152 и сл.; *Липкин, М. И*. Кривые распределения в экономических исследованиях. Москва, "Статистика",1972, с. 25; *Weisstein, E. W*. Pareto Distribution, http://mathworld.wolfram.com/ParetoDistribution.html; *Reed, W*. The Pareto, Zipf and other power laws. University of Victoria, Discussion Paper, 2000; *Souma, W*. Universal Structure of the Personal Income Distribution, arXiv:cond-mat/0011373 v.1 22, 2000 и др.

$$F(x/\alpha,c) = \int_X^{\infty} f(x)dx = 1 - \left(\frac{c}{x}\right)^{\alpha}$$, за всяко $x \geq c$,

където: $c > 0$ е най-малкото значение на променливата x_i, като единиците, притежаващи значение на признака под него, не се включват в съвкупността;

$\alpha > 0$ - параметър на уравнението.

Оттам функцията на плътността на разпределението на Парето има вида:

$$f(x/\alpha,c) = \frac{\alpha c^{\alpha}}{x^{\alpha+1}} = \frac{\alpha}{c}\left(\frac{c}{x}\right)^{\alpha+1}.$$

Както беше отбелязано, параметърът c характеризира най-ниската стойност на променливата, от която започва кривата на Парето. От уравнението на кривата се вижда, че когато $x \to \infty$, $f(x) \to 0$, а когато $x \to c$, $f(x) \to \infty$. Следователно кривата на Парето има две асимптоти: $x = c$ и $f(x) = 0$. На практика това означава, че значенията на признака не започват от най-малките му стойности, а от субективно определен размер, равен на c.

Параметърът α се приема като мярка за разпределението на масата на признака между единиците на съвкупността и е в основата на т. нар. принцип на Парето.[5] Когато разпределението на Парето се използва за моделиране, важно значение има съотношението между минималното значение на случайната променлива $x_{\min} = c$ и коя да е стойност на променливата x_i, като $x_i > x_{\min}$, т.е. $\left(\frac{c}{x}\right)^{\alpha}$. Тъй като съотношението трябва да бъде между 0 и 1, индексът α трябва да бъде положително число. Колкото е по-голямо значението на α, толкова са по-ниски вероятностите за големите значения на променливата. В този смисъл параметърът α може да се използва като мярка за неравномерността.

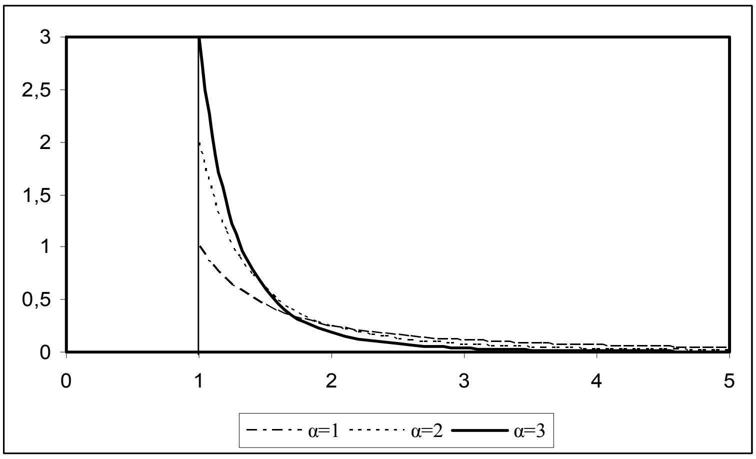
На фиг. 3. е представен графичният образ на кривата на Парето за различни значения на c и α. С нарастването на α кривата на разпределението става по-вдлъбната.

---

[5] Основен икономически принос на В. Парето е установеният от него “Принцип 80-20”, според който 20% от населението притежават 80% от доходите на цялото общество. Този принцип е доразвит от Дж. М. Джиран (J.M.J) - американски технолог и филантроп със съществен принос в теорията на управлението, който дава името му през 1941 г. в памет на В. Парето (вж. Pareto index. ISI glossary of statistical terms, http://epp.eurostat.ec.europa.eu.).

Фигура 3

Функция на плътността на разпределението на Парето при различни стойности на α



Отсеченото нормално разпределение[6] намира широко приложение в различни области на емпиричните изследвания, насочени преди всичко към оценка на параметрите на генерални съвкупности с нормално (или близко до нормалното) разпределение въз основа на данни от "отсечени" съвкупности. Обстоятелството, че характеристиките на нормално разпределена случайна величина могат да приемат всяка стойност в интервала от $-\infty$ до $+\infty$, може да доведе до съществени грешки в случаите, когато емпиричното разпределение е дефинирано в по-тесни граници.

Отсеченото разпределение е условно разпределение на случайна променлива. По правило то е получено от някакво друго разпределение и има най-общо следната функция на плътността:

$$f(x/x>y)=\frac{f(x)}{1-F(y)},$$

където: $f(x)$ е функцията на непрекъснатото разпределение, чиято дефиниционна област е предефинирана от $(-\infty;+\infty)$ на $(y,+\infty)$;

[6] *Съйкова, И., А. Стойкова-Къналиева и С. Съйкова*. Статистическо изследване на зависимости. С.: УИ "Стопанство", 2002, с. 359; *Madala, G. S*. Introduction to econometrics. New York: Macmillan, 1988, p. 283; *Johnson, A. C., N. T. Thomopoulos*. Characteristics and tables of the left-truncated normal distribution. Proceedings of Midwest Decision Sciences Institute, 2002, p. 133-139; *Johnson, A. C., N. T. Thomopoulos.* Use of the Left-Truncated Normal Distribution for Improving Achieved Service Levels. Proceedings of the Decision Sciences Institute, 2002, p. 2033-2041; Truncated distribution, http://en.wikipedia.org/wiki/Truncated_Distribution.

$F(x)$ - кумулативната функция на същото разпределение с нормална дефиниционна област, при условие, че и двете са известни.

В емпиричните статистически изследвания обикновено значенията на признака приемат стойности в ограничен интервал (както в малките значения на признака, така и в големите). Подобно на кривата на Парето отсеченото нормално разпределение характеризира разпределението на единиците в съвкупности, от които са отнети (не са включени) част от единиците. Това са онези единици, които имат значения на признака до определена стойност (y), но преди или след центъра на разпределението.

Ако съществуват значения на x_i, които са по-малки от точка x_L (точка на отсичане), които не са наблюдавани поради определени причини, то е налице ляво отсечено нормално разпределение с функция на плътността

$$f_L(x)=\begin{cases} 0, -\infty \le x \le \infty \\ \dfrac{f(x)}{\int_{x_L}^{\infty} f(x)dx} & , x_L \le x \le \infty, \end{cases}$$

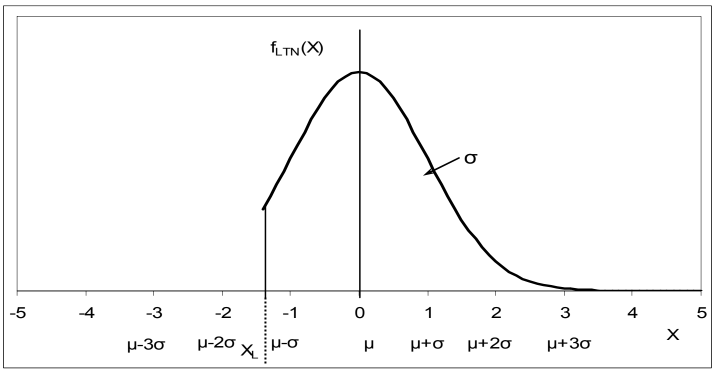
В контекста на стандартизираното нормално разпределение $f(z)$, където $z=\dfrac{x-\mu}{\sigma}$, функцията на плътността придобива вида:

$$f(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}, \ -\infty \le z \le \infty .$$

Графичният образ на кривата на разпределението е представен на фиг. 4.

Фигура 4

Функция на плътността на ляво отсечено нормално разпределение (емпиричните характеристики)



Ако точката на отсичане x_L се стандартизира с нейната оценка $K_L = \frac{x_L - \mu}{\sigma}$, началото на кривата на разпределението се измества в точка K_L. За да се стандартизира ляво отсеченото нормално разпределение в областта $(0,\infty)$, е необходимо да се въведе стандартната променлива $t = z - K_L$. Тя е получена като резултат от дефинирането на точката на отсичане при $t = 0$. Стандартизираното по този начин ляво отсечено нормално разпределение има функция на плътността $f_{SLTN}(t)$, определена като:

$$f_{SLTN}(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} .$$

Тази форма на връзка се използва при определяне на теоретичните честоти при моделиране на вариационното разпределение. Математическото очакване и дисперсията на стандартизираното ляво отсечено нормално разпределение за определена точка на отсичане K_L зависи единствено и само от точката на отсичане.

Оценка на параметрите и установяване на теоретичните честоти на разпределенията

За установяване на теоретичните честоти на разпределенията е необходимо да се извърши оценка на параметрите на трите функции на разпределение. В теорията са разработени няколко метода за точкови оцени на неизвестните параметри, но най-широко практическо приложение намират два от тях - методът на моментите и методът на максималното правдоподобие.[7]

Теоретичната функция на плътността на разпределението може да се изрази като $f\left(x / k_1, k_2, ..., k_p\right)$, която зависи от p на брой параметри: $k_1, k_2, ..., k_p$. За избраните от нас разпределения това са:

[7] За повече подробности вж. *Съйкова, И., А. Стойкова-Къналиева, С. Съйкова*. Цит. съч., с. 67 - 78; *Липкин, М. И.* Цит. съч., с. 70-76; *Spiegel, M. R*. Theory and Problems of Probability and Statistics. New York: McGraw-Hill, 1992; *Paul, T. K., I. Hitoshi*. Linear and Combinatorial Optimizations by Estimation of Distribution Algorithms. 9th MPS symposium on Evolutionary Computation, IPSJ, Japan, 2002 и др.

●за експоненциалното разпределение: $k_1 = \lambda$.

●за разпределението на Парето: $k_1 = \alpha$; $k_2 = c$.

●за ляво отсеченото нормално разпределение: $k_1 = \mu$; $k_2 = \sigma$; $k_3 = K_L$.

Получените по метода на моментите оценки са състоятелни, но невинаги ефективни и достатъчни, особено когато са изчислени по данни от извадки. Моментите, изчислени от извадки, са приблизителни оценки на моментите в генералните съвкупности. Когато оценките са получени въз основа на данни от изчерпателни наблюдения (какъвто е случаят в това изследване), оценяването спира до изчисляването на точковата оценка. Само ако те са от извадкови изследвания, е необходимо да се оцени и стохастичният компонент на грешката. Освен това оценката на параметрите при експоненциалното и нормалното разпределение (в частност ляво отсеченото нормално), установени по метода на моментите, съвпада с оценката на параметрите по метода на максималното правдоподобие. Именно поради тези обстоятелства е използван методът на моментите за оценка на неизвестните параметри на изследваните разпределения.

Оценка на параметъра λ на експоненциалното разпределение. Тъй като параметърът λ е единствен във функцията на разпределението, може да се очаква, че емпиричното и теоретичното значение на първия начален момент са равни. Така λ може да се изчисли въз основа на равенството между математическото очакване на теоретичното разпределение и емпиричната средна аритметична величина или средното квадратично отклонение. Това е възможно, ако емпиричните характеристики на $\bar{x}$ и $\sigma_{(x)}$ са приблизително равни. Резултатите от изчисляването на обобщаващите числови характеристики показват, че разсейването в средния размер на земята, измерено чрез средното квадратично отклонение, е в пъти по-голямо, поради което приемаме като база за оценка на λ средната аритметична величина, изчислена от груповите средни. На табл. 2. са представени стойностите на параметъра за отделните моменти на наблюдение.

Таблица 2

Оценки на параметъра λ на теоретичната крива на експоненциалното разпределение

Години	1897	1908	1926	1934	1946	2001	2003	2005
λ	0.02011	0.02018	0.01749	0.02026	0.02343	0.15337	0.02255	0.01907

Оценка на параметрите на разпределението на Парето. Параметърът C характеризира най-ниската граница на вариране на случайната променлива. Оценката на α се основава на предположението, че C е известно. В конкретното изследване това е най-малкият размер стопанисвана земя, притежавана от земеделско стопанство. Ако разполагаме с индивидуални данни, определянето на C не е свързано с никакви затруднения. При наличието на обобщени (групирани) данни са възможни няколко варианта за неговата оценка. Единият е за C да се приеме минималният размер земя, използван като "праг" за включване на земеделското стопанство като единица на наблюдение. През продължителния период, обхванат в изследването, този "праг" е бил твърде различен. Освен това има стопанства със земя под "прага", но включени като животновъдни стопанства. Третата възможност е да се избере модата, изчислена от групирани данни, но поради крайно асиметричната форма на разпределение трябва изкуствено да отпаднат голяма част от единиците със земя, реално по-малка от тази стойност. Поради това за оценка на C приехме груповата средна аритметична величина в първата група от 0 до 10 дка (вж. табл. 3.). При положение, че липсват индивидуалните данни за наблюдаваните единици, значенията на случайната променлива са средните групови в отделните групи по значенията на признака. Така средната групова в първата група е най-малката стойност на x_i, т.е. тя е най-близко до реалната стойност $x_{\min}$.

Таблица 3

Оценки на параметрите на теоретичната крива на разпределението на Парето

Години	α	C
1897	1.093038	4.233774
1908	1.098036	4.424851
1926	1.094701	4.945485
1934	1.109759	4.882677
1946	1.130073	4.913207
2001	1.447582	2.015936
2003	1.094701	4.945485
2005	1.076659	3.733361

Оценка на параметрите на ляво отсеченото нормално разпределение Оценката на параметрите на тази разновидност на нормалното разпределение не се различава принципно от оценката на параметрите на нормалното стандартно разпределение. Плътността на разпределението зависи от два параметъра – математическото очакване и дисперсията, респ. средното квадратично отклонение.

За крайно асиметричната форма на емпиричното разпределение приехме модата за по-подходящ център на разпределение от средната аритметична. Поради това и нормираните отклонения z_i са изчислени на базата на модата, а не на средната аритметична. Както вече отбелязахме, характеристиката, която определя спецификата на ляво отсеченото нормално разпределение, е точката на отсичане x_L и нейните нормирани (стандартизирани) оценки K_L. За точка на отсичане се приема средната групова от фактическите значения на първата група от 0 до 10 дка, основавайки се на аргументите за оценка на параметрите при кривата на Парето. При наличието на групирани данни средната групова в първата група е най-малкото значение на случайната променлива (табл. 4.).

Таблица 4

Оценки на параметрите на теоретичната крива на ляво отсеченото нормално разпределение

Години	$\hat{\mu}$	$\hat{\sigma}$	K_L
1897	6.302	96.802	-0.02136
1908	6.437	86.160	-0.02335
1926	15.576	56.096	-0.18951
1934	10.546	47.489	-0.22517
1946	10.251	40.755	-0.13097
2001	5.298	31.772	-0.10333
2003	5.468	394.67	-0.00399
2005	5.532	412.49	-0.11768

Изравняването на емпиричните честоти е извършено по уравненията на трите функции за всички моменти на наблюдение (вж. табл. 5).

Теоретичните честоти на разпределението по експоненциалната крива са представени на табл. 6, на разпределението по кривата на Парето – на табл. 7 и на разпределението по кривата на ляво отсеченото нормално разпределение - на табл. 8.

Вследствие от изглаждането се получават криви на разпределение, които отстраняват в различна степен отделните неравномерности в емпиричните редове, резултат от действието на случайни фактори. Тъй като при моделирането са използвани три функции на разпределение, логически възниква въпросът за съответствието между емпиричните разпределения и получените теоретични разпределения. За да се установи това, са възможни различни подходи в зависимост от изследователските задачи.

Таблица 5

Уравнения на теоретичните криви на разпределението на земеделските стопанства

Години	Експоненциално разпределение	Разпределение на Парето	Ляво отсечено нормално разпределение
1897	$f_i = 0{,}02011.e^{-0{,}02011x_i}$	$f_i = \frac{1{,}033038.4{,}233774^{1{,}033038}}{x_i^{1{,}033038+1}}$	$f_i = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left[\left(\frac{x_i-6{,}302}{96{,}802}\right)+0{,}02136\right]}$
1908	$f_i = 0{,}02018.e^{-0{,}02018x_i}$	$f_i = \frac{1{,}098036.4{,}424851^{1{,}098036}}{x_i^{1{,}098036+1}}$	$f_i = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left[\left(\frac{x_i-6{,}437}{86{,}160}\right)+0{,}02335\right]}$
1926	$f_i = 0{,}01749.e^{-0{,}001749x_i}$	$f_i = \frac{1{,}094701.4{,}945485^{1{,}094701}}{x_i^{1{,}094701+1}}$	$f_i = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left[\left(\frac{x_i-15{,}576}{56{,}096}\right)+0{,}18951\right]}$
1934	$f_i = 0{,}02026.e^{-0{,}02026x_i}$	$f_i = \frac{1{,}109759.4{,}882677^{1{,}109759}}{x_i^{1{,}109759+1}}$	$f_i = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left[\left(\frac{x_i-10{,}546}{47{,}489}\right)+0{,}22517\right]}$
1946	$f_i = 0{,}02343.e^{-0{,}02343x_i}$	$f_i = \frac{1{,}130073.4{,}913207^{1{,}130073}}{x_i^{1{,}130073+1}}$	$f_i = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left[\left(\frac{x_i-10{,}251}{40{,}755}\right)+0{,}13097\right]}$
2001	$f_i = 0{,}15337.e^{-0{,}15337x_i}$	$f_i = \frac{1{,}447582.2{,}015936^{1{,}447582}}{x_i^{1{,}447582+1}}$	$f_i = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left[\left(\frac{x_i-5{,}298}{31{,}772}\right)+0{,}10333\right]}$
2003	$f_i = 0{,}02255.e^{-0{,}02255x_i}$	$f_i = \frac{1{,}094701.4{,}945485^{1{,}094701}}{x_i^{1{,}094701+1}}$	$f_i = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left[\left(\frac{x_i-5{,}468}{394{,}67}\right)+0{,}00399\right]}$
2005	$f_i = 0{,}01907.e^{-0{,}01907x_i}$	$f_i = \frac{1{,}076659.3{,}733361^{1{,}076659}}{x_i^{1{,}076659+1}}$	$f_i = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left[\left(\frac{x_i-5{,}532}{412{,}49}\right)+0{,}11768\right]}$

Таблица 6

Теоретични честоти по функцията на експоненциалното разпределение

x_i (дка)	$f_i(x_i/\lambda)=\lambda e^{-\lambda x_i}$							
	1897 г.	1908 г.	1926 г.	1934 г.	1946 г.	2001 г.	2003 г.	2005 г.
До 10	145 635	170 546	120 458	162 252	217 009	1 142 377	132 167	90 381
Над 10 " 20	119 111	139 384	101 127	132 501	171 687	246 441	105 490	74 688
" 20 " 30	97 418	113 915	84 898	108 205	135 830	53 164	84 198	61 720
" 30 " 40	79 676	93 101	71 274	88 364	107 462	11 469	67 203	51 003
" 40 " 50	65 165	76 089	59 836	72 162	85 019	2474	53 639	42 147
" 50 " 60	53 297	62 186	50 233	58 930	67 263	534	42 812	34 829
" 60 " 70	43 590	50 823	42 172	48 124	53 215	115	34 171	28 782
" 70 " 80	35 651	41 537	35 404	39300	42 101	25	27 274	23 784
" 80 " 90	29 158	33 947	29 723	32 094	33 308	5	21 769	19 655
" 90 " 100	23 848	27 744	24 953	26 209	26 352	1	17 375	16 242
" 100 " 150	67 901	78 846	76 100	74 332	68 884	0	46 452	47 512
" 150 " 200	24 849	28 749	31 735	26 997	21 351	0	15 047	18 309
" 200 " 300	12 422	14 305	18 753	13 367	8669	0	6453	9775
" 300 " 400	1664	1902	3261	1763	833	0	677	1452
" 400 " 500	223	253	567	233	80	0	71	216
Над 500	34	39	119	35	9	0	8	38
Всичко	799 642	933 367	120 458	162 252	1 039 072	1 456 606	654 808	520 533

Таблица 7

Теоретични честоти по функцията на разпределението на Парето

x_i (дка)	$f_i(x_i/\alpha,c)=\frac{\alpha c^{\alpha}}{x_i^{\alpha+1}}$							
	1897 г.	1908 г.	1926 г.	1934 г.	1946 г.	2001 г.	2003 г.	2005 г.
До 10	487 110	552 094	403 344	485 509	573 630	1 313 217	351 863	340 337
Над 10 " 20	166 026	203 161	184 666	214 308	252 785	90 817	161 096	94 761
" 20 " 30	52 451	63 998	58 285	67 054	78 169	23340	50 845	30 222
" 30 " 40	25 377	30 909	28 182	32 250	37 327	9956	24 585	14 707
" 40 " 50	14 865	18 082	16 501	18 809	21 656	5321	14 395	8651
" 50 " 60	9723	11 815	10 790	12 262	14 059	3237	9413	5678
" 60 " 70	6837	8301	7585	8597	9824	2143	6617	4003
" 70 " 80	5059	6138	5611	6347	7231	1507	4895	2969
" 80 " 90	3889	4716	4313	4868	5532	1108	3762	2287
" 90 " 100	3079	3731	3414	3847	4362	843	2978	1814
" 100 " 150	9031	10 931	10 009	11 239	12 681	2271	8731	5343
" 150 " 200	4370	5279	4840	5406	6055	969	4222	2600
" 200 " 300	4234	5107	4687	5208	5794	833	4088	2533
" 300 " 400	2048	2466	2266	2505	2767	355	1977	1233
" 400 " 500	1200	1443	1327	1461	1605	190	1157	725
Над 500	4344	5196	4795	5199	5596	498	4183	2670
Всичко	799 642	933 367	750 613	884 869	573 630	1 456 606	654 808	520 533

Таблица 8

Теоретични честоти по функцията на ляво отсеченото нормално разпределение

x_i (дка)	$f_i(x_i / \mu,\sigma,K_L) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-K_L)}$							
	1897 г.	1908 г.	1926 г.	1934 г.	1946 г.	2001 г.	2003 г.	2005 г.
До 10	61 608	80 192	85 035	124 172	166 554	307 897	16 408	12 883
Над 10 “ 20	61 389	79 854	86 564	124 606	166 988	300 126	16 404	12 881
“ 20 “ 30	60 490	78 403	85 464	119 715	157 643	263 955	16 390	12 870
“ 30 “ 40	58 986	75 955	81 755	110 147	140 417	211 113	16 365	12 852
“ 40 “ 50	56 901	72 594	75 742	97 132	117 770	152 458	16 328	12 825
“ 50 “ 60	52 344	65 336	68 106	81 716	93 284	101 681	16 283	12 796
“ 60 “ 70	52 344	65 336	59 150	65 745	69 334	60 030	16 227	12 752
“ 70 “ 80	49 281	60 408	49 784	50 483	48 508	31 939	16 160	12 704
“ 80 “ 90	43 743	51 999	40 687	37 290	32 341	15 881	16 085	12 653
“ 90 “ 100	43 743	51 999	32 081	26 197	20 246	6715	15 986	12 570
“ 100 “ 150	152 572	165 471	76 114	45 313	25 504	4805	78 703	62 027
“ 150 “ 200	71 715	63 959	9611	2225	462	6	75 057	59 313
“ 200 “ 300	33 007	21 449	394	89	12	0	137 005	109 351
“ 300 “ 400	1497	409	113	19	7	0	114 013	90 949
“ 400 “ 500	21	2	12	17	2	0	87 393	71 108
Над 500	2	0	1	3	0	0	0	0
Всичко	799 642	933 367	750 613	884 869	1 039 072	1 456 606	654 808	520 533

Оценка на съответствието между емпиричните и теоретичните разпределения

В теорията на статистиката са разработени редица обективни методи за оценка на съответствието между емпиричните и теоретичните разпределения, т. нар. критерии на съгласие. Най-широко приложение намира в подобни случаи χ^2 критерият като мярка за различията между емпиричните и теоретичните честоти.[8] Известно е, че при съвпадение на емпиричните и теоретичните криви $\chi^2 = 0$. Горната граница на величината на критерия клони към $+\infty$, т.е. няма ограничения във варирането му. Оценката на χ^2 критерия има вероятностен характер, като вероятността намалява от $1 \to 0$, когато значението на χ^2 нараства от $0 \to +\infty$. Табличната стойност на χ^2 показва границата,

[8] *Съйкова, И., А. Стойкова-Къналиева, С. Съйкова*. Цит. съч., с. 75 - 78; *Петров, В., П. Ангелова, К. Славева*. Статистически методи за изследвания в социалната сфера. - Стопански свят, АИ “Д. А. Ценов”, Свищов, 2006, N 78, с. 83-90; *Дружинин, М. К.* Математическая статистика в экономике. Москва, “Статистика”, 1971, с. 243 и сл.

превишението на която при известни степени на свобода свидетелства за наличието на съществени отклонения. Оценяването чрез χ^2 критерия зависи освен от степените на свобода и от самата величина на този показател. В теорията са известни изискванията за неговото използване, свързани с ограничения по отношение на минималния брой единици в наблюдаваната съвкупност; минимален брой единици в отделните групи; минимален брой групи и др. Не са конкретно дефинирани обаче условията за приложение при максимизиране на тези характеристики. Използването на χ^2 критерия при големи по обем извадки и много на брой групи води до получаването на многоцифрени стойности на неговата емпирична величина. Така нейното сравнение с табличната му стойност става некоректно и винаги свидетелства за съществено разминаване между емпиричното и теоретичното разпределение.

В конкретното изследване моделите на теоретичните криви се основават на оценки на параметрите в изчерпателни съвкупности, които са с много голям обем. Ето защо за някои наблюдения се получиха шестцифрени стойности за χ^2, които не могат да се сравняват с табличните. Именно поради това χ^2 критерият не може да се приложи в случая, тъй като се изравнява разпределението на изчерпателна съвкупност, а не в извадка.[9]

Този съществен аргумент ни дава основание да приемем друг подход за оценяване на съответствието между емпиричните и теоретичните криви, а именно да се оцени коя от избраните теоретични функции най-близо, най-добре възпроизвежда емпиричното разпределение за даден момент.

Работните хипотези, които дефинираме при моделиране на разпределението, се изразяват в следното:

- За 1897 и 1908 г. най-близко до формата на емпиричното разпределение е кривата на експоненциалното разпределение.
- За 1926, 1934 и 1946 г. най-близко до формата на емпиричното разпределение е кривата на ляво отсеченото нормално разпределение.
- За 2001, 2003 и 2005 г. най-близко до формата на емпиричното разпределение е кривата на Парето.

За проверка на съответствието използваме два от известните показателя за минимизиране на различията между емпирични и изгладени значения:

- За сравнение на абсолютните емпирични и теоретични честоти прилагаме стандартната грешка на оценката, адаптирана по следния начин:

$$S_f = \sqrt{\frac{\sum\left(f_i - \hat{f}_i\right)^2}{k}} ,$$

[9] *Липкин, М. И.* Цит. съч., с. 103.

където: f_i е емпиричната честота в съответната група;

$\hat{f}_i$ - теоретичната честота в същата група;

k - броят на групите.

• За сравнение на относителните емпирични и теоретични честоти прилагаме интегралния коефициент на структурни различия, модифициран в случая:

$$K_S = \sqrt{\frac{\sum\left(V_i - \hat{V}_i\right)^2}{\sum V_i^2 + \sum \hat{V}_i^2}},$$

където: V_i е емпиричният относителен дял на единиците в групата;

$\hat{V}_i$ - теоретичният относителен дял на единиците в същата група.

Тъй като интегралният коефициент е стандартизиран в границите от 0 до 1, това го прави в случая и подходящ за оценка на степента на доближаване на теоретичната крива до емпиричната.

Така работните хипотези се свеждат до проверка на близостта на теоретичната функция до съответното емпирично разпределение. Още повече, че целта на моделирането не е да се открие нова функция, точно възпроизвеждаща закономерността в емпиричното разпределение, а да се определи до кое от известните, вече установени в теорията закони на разпределението, е най-близо емпиричното разпределение на земеделските стопанства според размера на земята.

На табл. 9 и 10 са представени величините на стандартната грешка и коефициента на структурни различия между емпиричните и теоретичните честоти на разпределението.

Таблица 9

Стандартна грешка на оценката $\left(S_f\right)$

Години	Експоненциално разпределение	Разпределение на Парето	Ляво отсечено нормално разпределение
1897	*29 022*	61 795	58 398
1908	*32 222*	85 875	62 516
1926	10 050	85 875	*7145*
1934	14 270	99 495	*7717*
1946	20 502	114 231	*9097*
2001	31 768	*28 299*	244 276
2003	95 626	*85 875*	136 071
2005	75 453	*12 365*	132 624

Таблица 10

Коефициент на структурни различия (K_S)

Години	Експоненциално разпределение	Разпределение на Парето	Ляво отсечено нормално разпределение
1897	*0.043112*	0.088547	0.173813
1908	*0.039361*	0.086188	0.148601
1926	0.008026	0.246951	*0.003859*
1934	0.010580	0.228927	*0.002904*
1946	0.014234	0.212244	*0.002617*
2001	0.003004	*0.002114*	0.261617
2003	0.246998	*0.036457*	0.465005
2005	0.264748	*0.004625*	0.465452

От посочените в таблиците значения се установява следното:

●За първите две наблюдения най-близко до емпиричното разпределение е кривата на експоненциалното разпределение с минимални величини на стандартната грешка $S_{1897} = 29022$ и $S_{1908} = 32222$, съответно с минимални стойности на коефициента на структурни различия $K_{S1897} = 0{,}043112$ и $K_{S1908} = 0{,}039361.$

●За разпределенията през 1926, 1934 и 1946 г. най-близо до емпиричната форма е изгладената по функцията на ляво отсеченото нормално разпределение крива. Минималните стойности на стандартната грешка са: $S_{1926} = 7145$, $S_{1934} = 7717$ и $S_{1926} = 9087$, като за другите криви разликата е няколко пъти по-голяма и например през 1946 г. достига по Парето до 114231. Същото се отнася и за коефициента на структурни различия, който приема освен минималните спрямо другите криви значения, но и изключително ниски спрямо нормираните му граници стойности, съответно $K_{S1926} = 0{,}003859$, $K_{S1934} = 0{,}002904$ и $K_{S1946} = 0{,}002617$.

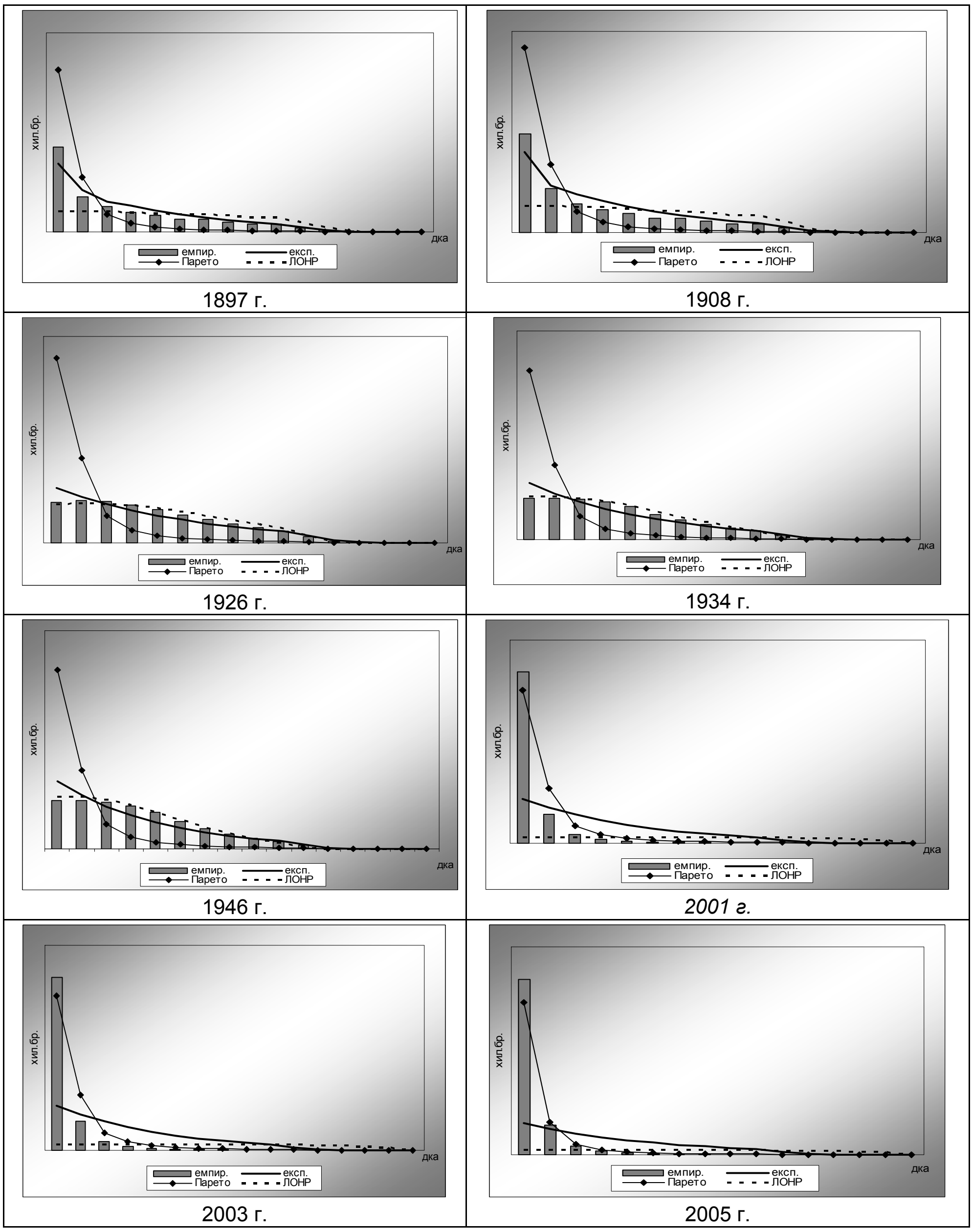
●За последните три наблюдения през 2001, 2003 и 2005 г. най-добро изглаждане се получава по кривата на Парето. Стандартната грешка има минимални стойности за тези моменти, които са съответно $S_{2001} = 28299$, $S_{2003} = 85875$ и $S_{2005} = 12365$. Относителните различия, измерени чрез коефициента на структурни изменения, също приемат минимални стойности, възлизащи на: $K_{S2001} = 0{,}002114$, $K_{S2003} = 0{,}036457$ и $K_{S2005} = 0{,}004625$.

Представените на фиг. 5 графични образи на емпиричните и теоретични криви на разпределението потвърждават направените изводи. [10]

[10] За техническа прегледност използвахме следните съкращения в легендата на графичните изображения: емпир. = емпирична крива; експ.=експоненциална крива; Парето = крива на Парето и ЛОНР = крива на ляво отсеченото нормално разпределение.

Фигура 5

Емпирични и теоретични криви на разпределението



*

Изследването на динамиката на вариационните разпределения дава възможност да се решат две изследователски задачи. Изглаждането на колебанията в емпиричното разпределение, предизвикани от влиянието на променливи във времето фактори, позволява да се установят неговите основни, типични черти към конкретен момент, в т.ч. и на функцията на разпределението. Същевременно може да се открои еволюцията на самата форма на разпределение, довела до постепенно качествено изменение във вида на функцията (интегрална и диференциална) на разпределението. В резултат от конкретното изследване се установи, че динамиката на разпределението на земеделските стопанства според размера на стопанисваната земя се характеризира с преход от една закономерност към друга. Този преход може да се обясни със сериозните промени в материалните условия, определящи състоянието на българското земеделие през различните етапи от неговото развитие.

Отчитайки конкретните особености на разпределенията в изследваните съвкупности и резултатите от моделирането на функциите на плътността на разпределението и нейното изменение през периода на изследване, можем да направим следните изводи:

- През разглеждания период не се откроява точно определена единна функция, описваща закономерността на разпределение. Независимо от това разпределението на земеделските стопанства се характеризира със специфична крайно асиметрична L-образна форма, като през отделните етапи на развитие кривата има повече или по-малко вдлъбната форма, т.е. различна степен на неравномерност на разпределение на единиците по групи.
- Използваните при моделирането функции на разпределение са близки по своята форма и същност. Експоненциалното разпределение и разпределението на Парето са от групата на показателните разпределения, а ляво отсеченото нормално разпределение и разпределението на Парето характеризират закономерностите в т.нар. отсечени съвкупности (разликата е в мястото на точката на отсичане – преди или от модалното значение). Поради това в различните моменти на наблюдение емпиричното разпределение се доближава и може сравнително добре да се апроксимира с функцията на едно от трите разпределения.
- Динамиката на формата на разпределение се характеризира с преход от една функция на плътността към друга. При това на промеждутъчния етап от изследвания период формата на разпределение не съответства нито на изходния етап (експоненциална форма), нито на следващия (кривата на Парето). Но именно през този етап на развитие формата се доближава до кривата на нормалното разпределение, макар и със силно изразена асиметрия.

Крайно асиметричната форма на разпределение на земеделските стопанства (на съвременния етап от развитието на селското стопанство) свидетелства за задълбочаване на неравномерността в разпределението на земята и връщане на поземлената структура към предвоенната 1939 г. Разпокъсаната поземлена собственост и нерационалното разпределение на поземлените участъци възпрепятстват ефективното управление и използване на земята. Тесните рамки на дребните частни стопанства принуждават земеделските собственици да поддържат екстензивна експлоатация на земите си и да търсят пътища за засилване жизнеспособността на стопанствата.

В икономически и организационно-стопански аспект е необходимо да се запазят и укрепят целостта и размерът на възстановените и новосъздадените земеделски стопанства. Като решаваща предпоставка за нарастване на икономическия и производствения им потенциал се смята преди всичко увеличаването на размера на стопанствата. Възможността да се увеличи размера на стопанисваната земя трябва да се използва във всички категории стопанства независимо от тяхната организационна форма. Успехът на поземленото реформиране зависи от повишаването на продуктивността на разпределените или преразпределените земеделски земи. Когато това нарастване е под съмнение поради дребния размер на поземлените участъци, се налага да се потърсят възможностите на стопанските форми от корпоративен тип. Особеностите на пазарното земеделие предполагат масовото използване на такива земеделски стопанства, тъй като те осигуряват по-ефективно производство, без да се сменя собственикът на земята.

Присъединяването на България към ЕС предполага, че ще е необходимо да се използва опитът на европейските държави в структурната селскостопанска политика. Независимо от националните и географските особености на страните през различните периоди от развитието на Общността, по отношение на разпределението на земеделските стопанства според размера на земята е провеждана диференцирана политика, главно в две направления. От една страна, се използват мерки и инструменти за уедряване на земеделските стопанства, а от друга, се въздейства за укрепване на дребните земеделски стопанства в райони с проблеми от демографски и икономически характер и засилване на тяхната жизнеспособност. Типизирането на европейските земеделски стопанства на едри, средни и дребни по своето съдържание се различава от използваните в България групи, тъй като са различни границите на групите според размера на земята. Независимо от това, макар и в други мащаби, реформирането на разпределението на земеделските стопанства трябва да бъде насочено към уедряване на земята и установяване на такива организационни структури, чрез които да се осигурят условия за модерно и ефективно земеделие.

3.VI.2009 г.