

---

## ПЪРВА ГЛАВА

# СТАТИСТИЧЕСКИЯТ АНАЛИЗ В СОЦИОЛОГИЧЕСКИ, ИКОНОМИЧЕСКИ И БИЗНЕС ИЗСЛЕДВАНИЯ

---

## 1. СТАТИСТИЧЕСКИЯТ ПОДХОД ЗА АНАЛИЗ: ХАРАКТЕРНИ ОСОБЕНОСТИ

### 1.1. Статистиката като наука – основни понятия

Книгата има за предмет избрани основни аспекти на приложението на статистическия анализ за целите на икономически, социологически или бизнес изследвания. За ефективното усвояване на изложения материал е необходимо запознаване с особеностите на статистиката като наука и практическа дейност. Повече детайли относно тази проблематика читателят може да намери в специализираната статистическа литература, доколкото настоящата книга не съдържа детайлно изложение в тази насока.<sup>1</sup>

Статистиката е наука, предоставяща систематизирани знания относно подходите и методите за събиране, обработка и анализ на *данни, отнасящи се за съвкупности*. **Статистическата съвкупност** е първично понятие в статистическата наука и следва да се осмисли интуитивно. Тя представлява една общност от единици, наречени "статистически", които се определят чрез т.нар. "дефиниционен признак". Този признак ясно и еднозначно указва коя единица е част от дадена изследвана съвкупност. Общият брой на единиците в една статистическа съвкупност се нарича "обем на съвкупността".

Например съвкупността на населението на област Пловдив към 31.12.2016 г. се определя чрез прилагането на следния дефиниционен признак: *"всяко физическо лице, което има постоянна адресна регистрация в населено място от административната територия на област Пловдив към 31.12.2016 г."* Така съвкупността на населението в област Пловдив се отграничава ясно от съвкупността на населението например в област Стара Загора към същата дата. По данни на Националния статистически институт (НСИ) обемът на съвкупността на населението в област Пловдив към 31.12.2016 г. възлиза на 671 573 лица. Статистическите единици, формиращи съвкупността на активните предприятия в отрасъл "Хотелиерство и ресторантьорство" през 2015 г., се определят чрез дефиниционния признак: *"всяко регистрирано в страната предприятие с основна дейност в икономически сектор "Хотелиерство и ресторантьорство", което е осъществявало стопанска дейност през 2015 г."* По данни на НСИ тази съвкупност е с обем 27 306 предприятия.

---

<sup>1</sup> Съйкова, Ив., А. Стойкова-Къналиева, Св. Съйкова. Статистическо изследване на зависимости. С., Унив. изд. "Стопанство", 2002; Манов, А. Многомерни статистически методи. С., Унив. изд. "Стопанство", 2002; Чипева, С. Статистически анализ на категорийни данни с SPSS. С., Унив. изд. "Стопанство", 2005; Павлова, В., С. Чипева. Статистика. С., изд. "Нова звезда", 2012; Калоянов, Т., В. Петров. Статистика. С., ИК на УНСС, 2014.

**Статистическият признак** представлява някакво свойство, качество, характеристика на единиците на една съвкупност. Всички тези единици са еднакви по своя **дефиниционен признак**, но могат да се различават по един или повече **изучавани (изследвани) признаци**. Те се делят на две основни групи, наречени още "типове" признаци – количествени и качествени. Първият тип се наричат още "метрирани" (или "вариационни"), а вторият тип съответно "неметрирани" (или "категорийни"). Наименованието на самите типове подсказва по какво те се различават, а именно възможността техните значения да се измерват по някакъв количествен способ.

Например всяка единица на съвкупността на населението на област Пловдив, бидейки дадено физическо лице, притежава определено значение по всеки от следните признаци:

- пол: "мъж" или "жена";
- възраст: 26 г., 35 г., 2 г., 72 г. или някакъв друг брой навършени години;
- ръст: 185 см, 162 см, 174 см или някаква друга височина, измерена в сантиметри;
- семейно положение: лицето може да бъде "несемейно" (неженен, неомъжена), "семейно" (женен, омъжена), "разведено" или "овдовяло";
- тегло: 65,3 кг; 84,8 кг; 48,7 кг или някаква друга тежест, измерена в килограми;
- икономическа активност: "заето лице", "безработно лице", "пенсионер", "учащ", или друга специфична категория икономически неактивно лице;
- брой деца: 0 деца, 1 дете, 2 деца, 3 деца и т.н.
- и други разнообразни признаци.

Количествените признаци биват два вида: прекъснати и непрекъснати (дискретни). В общия случай количествените признаци са непрекъснати – такива, които могат да получават множество различни значения в някакъв числов интервал. Прекъснатите признаци са особен случай на метрирани характеристики – те могат да приемат само определени, фиксирани значения (най-често цели числа) в даден интервал. Обичайно най-простият способ на измерване – броенето на някакви неделими обекти – на практика води до дефинирането на прекъснат количествен признак.

Характеристиките "възраст", "ръст", "тегло" и "брой деца" на физическите лица от примера са количествени признаци, тъй като техният характер допуска получаването на числово измерими значения. От тях типичен прекъснат (дискретен) признак е "брой деца" – измервателният процес тук строго дефинира възможността при дадена единица (физическо лице) да се измери или стойност 0, или 1, или 2, или 3 и т.н. – един сравнително ограничен диапазон от цели числа като възможни стойности. Освен това характерът на признака е такъв, че не допуска измерване като "0,5 броя деца", "1,7 броя деца", "2,43 броя деца" и др. под. Поради това в практиката определянето на един количествен признак като "дискретен" зависи от реалния смисъл и възможност неговите значения да се измерят количествено.

За всеки тип статистически признак са налице **измервателни скали**, на които се разполагат неговите значения. При качествените (неметрираните) признаци се използват три основни вида скали: номинална (nominal), ординална (ordinal) и рангова (rank). Номиналната скала съдържа набор от възможни значения на признака, които са словесно определени и за които не е характерна никаква логическа подредба. Ординалната скала се различава от номиналната само по това, че възможните значения на признака следват определена последователност. Същевременно разликата между две нива по тази скала **няма количествена мярка!** Ранговата скала има специфичен подход за измерване – чрез нея се определя пореден номер (ранг) на всяка единица от съвкупността, получен чрез подреждане (ранжиране) на единиците по определен способ. Ранговата скала много често се разглежда като разновидност на ординалната скала.

Характеристиките "семейно положение" и "икономическа активност" на лицата са типичен пример за качествени статистически признаци, чиито значения са разположени

на номинална скала. Това е така поради факта, че не се вменява логическа последователност между категориите "несемейно", "семейно", "разведено" и "овдовяло" лице. Аналогичен е случаят с категориите "заето лице", "безработно лице", "пенсионер", "учащ" и т.н.

Признакът "пол" също е разположен на номинална скала, която обаче е от особен тип – нарича се "дихотомна" (или още "бинарна"). Самият вид качествен признак се определя още като "дихотомен" или "бинарен" поради факта, че може да получава само две възможни – взаимно изключващи се – значения ("мъж" или "жена").

За разлика от горепосочените признакът "степен на владеене на чужд език" е разположен на ординална скала. Неговите значения са подредени в следната логическа последователност, например: "не владее чужд език", "начинаещ", "средно напреднал", "напреднал" и "перфектно владеещ". По европейската система Europass нивата на владеене на чужд език се измерват чрез категоризиране по ординална скала с 6 последователни нива: A1, A2, B1, B2, C1 и C2. Степента на завършено образование на лицата също се измерва чрез признак, представен на ординална скала – неговите възможни значения са в следната наредба: "без образование", "начално", "основно", "средно", "висше – бакалавър", "висше – магистър", "висше – доктор". Характеристиката "степен на одобрение на данъчната политика" може да се измери чрез допитване до собствениците на фирми с помощта на въпрос, чиито възможни отговори са представени на специфичен вид ординална скала: "категорично не", "донякъде не", "нито не, нито да", "донякъде да" и "категорично да".

На рангова скала са разположени единиците в т.нар. "ранглисти" – например подредбата на тенисистите в класацията на АТП; подредбата на отборите в края на първенството по даден спорт и др. под. Ранжиране се прилага и в случаите, когато всяка единица бива поставена на първо, второ, трето и т.н. място от гледна точка на някакъв критерий – например представител на група експерти поставя пореден номер (ранг) на всяка от 6 застрахователни компании според личната му преценка за нивото на добронамереност на компанията при оценяването на щети: 1-ЗК "Омега", 2-ЗПК "Каппа", 3-ЗК "Бета", 4-ЗПК "Алфа", 5-ЗПК "Делта", 6-ЗК "Гама".

Количествените по характер признаци се измерват чрез метрирана скала, определяна още като "интервална".<sup>1</sup> В системата за управление на данните на използвания в следващите глави софтуер IBM SPSS Statistics, наричан по-нататък накратко SPSS, тя се означава като Scale. Метрираната скала позволява количествени съпоставки между нейни стойности, което дава възможност за провеждане на по-задълбочен статистически анализ чрез по-прецизни методи. Например непрекъснатият количествен признак "работна заплата" на единиците от съвкупността на наетите лица (измерена в лева) е представен на такава скала. Работната заплата 1260 лв. се съотнася към заплата от 630 лв. два пъти, при което 1260 лв. представляват два пъти по-голямо количество пари от 630 лв. Също така повишението на заплатата от 630 лв. до 830 лв. – съставляващо интервал от 200 лв. – е същото количество пари, с което се повишава заплатата от 1260 лв. до 1460 лв. Операциите в отрицателната страна на числовата ос имат аналогичен смисъл: намаление на дълг (или загуба) от –550 лв. до –450 лв. представлява повишение на стойността чрез добавяне на +100 лв. – и то е същото количество, което е намалението на дълга от –5550 лв. до –5450 лв. Същевременно дълг от –400 лв. е два пъти по-голям от дълг в размер на –200 лв.

<sup>1</sup> В редица източници по теория на статистиката се идентифицира т.нар. "относителна" (ratio) скала, при която частното на две стойности от нея отразява количественото съотношение между тези стойности. В рамката на настоящата книга е възприет традиционният подход за определяне на дадена метрирана скала като "интервална", прилагана при количествено измеримите признаци.