

# Глава 12

## Случайни извадки и статистика

Вече са ни известни някои дискретни и непрекъснати вероятностни разпределения като модели на реални ситуации, свързани с явления, които имат случаен характер. Един от естествените въпроси, които възникват е как да ги използваме на практика? Обикновено можем да решим какъв тип разпределение може да служи за модел на даден признак в определена ситуация, обаче параметрите които го определят точно не са известни. Например, при проучване на общественото мнение относно гласуване за общински съветници изборът е между “да” и “не”, следователно ще имаме биномно разпределение, но параметърът  $p$  не е известен; добивът на декар от селскостопанска култура е известно, че е нормална сл. в., но средното  $\mu$  и стандартното отклонение  $\sigma$  не са известни.

В такива случаи разчитаме на извадката да ни даде информация за точните стойности на тези неизвестни популационни параметри. В горните примери - процентът на тези, които отговарят с “да” ще ни даде представа за  $p$  и подобно извадковото средно и средно-квадратичното отклонение за  $\mu$  и  $\sigma$  на популацията. И в двата случая извадковата информация би била полезна да се направят изводи за популационните параметри.

Тъй като стойностите на извадковите статистики се менят от извадка на извадка и са непредсказуеми, то те са сл. в. Следователно, за да оценим надеждността на извода, базиран на извадковите стойности трябва да определим или апроксимираме извадковото разпределение на извадковата статистика.

**Определение 12.1** *Извадково разпределение на една статистика е вероятностното разпределение на стойностите на статистиката, които се получават, когато случайните извадки с обем  $n$  от популацията се повтарят многократно.*

Извадковото разпределение може да се получи с математически средства или да бъде емпирично.

**Определение 12.2** *Хистограма на относителните честоти наричаме графика, построена по наблюденията, където по оста  $x$  са нанесени интервалите, в които попадат стойностите на наблюденията, а по оста  $y$  са самите относителните честоти на тези наблюдения. Построената по този начин графика представлява стъпаловидна крива такава, че лицето което загражда с оста  $x$  трябва да бъде равно на единица.*

Хистограмата на данните ни дава представа за вероятностната плътност на разпределението.

Емпиричното разпределение се намира като се направят голям брой извадки с обем  $n$  от популацията, изчисли се стойността на статистиката за всяка извадка и се построи хистограмата на относителните честоти. Колкото броя на извадките е по-голям, толкова повече хистограмата се доближава до теоретичното извадково разпределение.

Знанието ни за извадковото разпределение на една статистика ни позволява да изберем най-добрата от всички кандидати за оценка на един неизвестен параметър. И също така, извадковото разпределение позволява да се оцени надеждността на използваната статистика чрез определяне на границите, в които очакваме да попадне оценката с определена вероятност.