

Az alapsokaság *várható értékét* a *minta átlagával*:

$$\mu \approx \bar{x}$$

szórását a *korrigált empirikus szórással*:

$$\sigma \approx s$$

a *valószínűséget* a *relatív gyakorisággal* becsüljük:

$$p \approx \frac{f}{n}$$

4.3.2. Intervallumbecslés

Igen kicsi annak a valószínűsége, hogy a minta alapján az alapsokaság ismeretlen α paraméterére számított érték megegyezzen a tényleges értékkel, hiszen a becsült érték mint valószínűségi változó kisebb-nagyobb ingadozást mutat a tényleges érték körül.

Annak érdekében, hogy képet kapjunk az elkövetett pontatlanságról, célszerű egy intervallumot megadni, amely nagy valószínűséggel tartalmazza az ismeretlen paramétert.

Definíció: Az $[\alpha_1; \alpha_2]$ véletlen helyzetű intervallumot az α paraméter becslésére szolgáló $(1 - \varepsilon)100\%$ *megbízhatósági szintű konfidenciaintervallumnak* nevezzük, ha a

$$P(\alpha_1 < \alpha < \alpha_2) = 1 - \varepsilon \quad \text{egyenlőség fennáll.}$$

Egy konkrét mintából kiszámított konfidencia intervallum vagy tartalmazza a tényleges paraméter értéket, vagy nem. Az $(1 - \varepsilon)$ megbízhatósági szint tehát azt jelent, hogy ha sokszor ismételnénk meg a mintavételt és mindegyik esetben kiszámítanánk a megbízhatósági intervallumot, akkor az esetek $(1 - \varepsilon)100\%$ -ában ezek az intervallumok tartalmazzák az alapsokaság ismeretlen paraméterét és csak $\varepsilon 100\%$ -ában nem.

A minta elemszámának növelésével adott $(1 - \varepsilon)$ -hoz tartozó konfidencia intervallum csökken, míg adott minta elemszám esetén ε csökkenésével egyre nő.

Minden paraméterre megadható a fentiekben definiált konfidencia intervallum. Mi most csak mintaként a leggyakrabban alkalmazottat adjuk meg.

Normális eloszlású alapsokaság várható értékének konfidencia intervalluma

$$\bar{x} - t_{1-\frac{\varepsilon}{2}} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{1-\frac{\varepsilon}{2}} \frac{s}{\sqrt{n}}$$