A bootstrap

Legyen $\xi_1, \ldots, \xi_n$ független minta az $f(x, \theta)$ sűrűségfüggvényű populációból. A bootstrap módszerrel a populáció θ paramétere $g(\theta)$ függvénye pontbecslésének szórását becsülhetjük és a $g(\theta)$ értékére konfidencia intervallumot adhatunk szimulációt alkalmazva.

Tekintsük a $T = t(\xi_1, \ldots, \xi_n)$ valószínűségi változót, $g(\theta)$ pontbecslését. Ha a sűrűségfüggvény alakja ismert, a pontbecslést felhasználva a sűrűségfüggvény adott, véletlenszám-generátorral ennek megfelelő eloszlású $\xi_{1b}, \ldots, \xi_{nb}$ mintát szimulálunk. Ha a sűrűségfüggvény alakja nem ismert, húzzunk visszatevéssel n elemű $\xi_{1b}, \ldots, \xi_{nb}$ véletlen mintát véletlenszám-generátorral az eredeti mintából. Számítsuk ki a $T_{1b} = t(\xi_{1b}, \ldots, \xi_{nb})$ értéket. A szimulációt ismételjük meg N-szer, a kapott értékek $T_{1b}, \ldots, T_{Nb}$. A bootstrap szórásnégyzet

$$S_{Nb}^2 = \frac{1}{N} \sum_{i=1}^{N} (T_{ib} - \overline{T_b})^2.$$

A $g(\theta)$ érték közelítő $100c\%$-os szintű konfidencia intervallumára az alábbi két módszert tekintjük:

a) normál konfidencia intervallum ha a minta normáleloszlású populációból származik: $(T - u_{(1+c)/2} S_{Nb}, T + u_{(1+c)/2} S_{Nb})$, ahol $\Phi(u_{(1+c)/2}) = (1 + c)/2$;

b) percentilis konfidencia intervallum: $(T - r_{(1+c)/2\,b}, T - r_{(1-c)/2\,b})$, ahol $r_{(1-c)/2\,b}$ a $T_{ib} - \overline{T_b}$ bootstrap minta $100(1 - c)/2$-edik percentilise, $r_{(1+c)/2\,b}$ a $T_{1b} - \overline{T_b}$ bootstrap minta $100((1 + c)/2)$-edik percentilise.