

POSZTEREK

Data-driven confidence intervals for the mean of heavy-tailed distributions

AMBRUS TAMÁS, SZENTPÉTERI SZABOLCS, CSÁJI BALÁZS CSANÁD

¹HUN-REN SZTAKI,

²Eötvös Loránd Tudományegyetem, Matematikai Intézet,
Valószínűségelméleti és Statisztika Tanszék

Estimating the expected value is one of the key problems of statistics, and it serves as a backbone for countless methods in machine learning. In this poster we present a new algorithm to build non-asymptotically exact confidence intervals for the mean of a symmetric distribution based on an independent, identically distributed sample. The method combines resampling with median-of-means estimates to ensure optimal subgaussian bounds for the sizes of the confidence intervals under mild, heavy-tailed moment conditions. The scheme is completely data-driven: the construction does not need any information about the moments, yet it manages to build exact confidence regions which shrink at the optimal rate.

Járványterjedés fertőzési rátájának becslése kétrétegű véletlen gráfokon

BACKHAUSZ ÁGNES, BOGNÁR EDIT, CSISZÁR VILLŐ,
TÁRKÁNYI DAMJÁN, ZEMPLÉNI ANDRÁS

Eötvös Loránd Tudományegyetem, Természettudományi Kar,
Valószínűségelméleti és Statisztika Tanszék

A járványterjedés modellezésénél kulcsfontosságú paraméter a fertőzési ráta. Ennek becslésére több módszert is bemutatunk: a klasszikus maximum likelihood alapú megközelítés mellett két olyan módszert is vizsgálunk, melyek gépi tanulást használnak, az XGBoostot, illetve konvolúciós neurális hálókat. A jól ismert SIR járványmodellt kétrétegű hálózatokon szimuláljuk. A hálózat első rétege kis teljes részgráfokból áll, melyek a háztartásoknak felelnek meg, míg a második – kisebb élsúlyokkal rendelkező – réteg az egyéb kapcsolatokat (pl. munkahelyek) reprezentálja. Kutatásunkban ez a második réteg az első rétegtől független, és vagy skálafüggetlen, vagy a háztartások rétegéhez hasonlóan fix méretű klikkekből épül fel.