Valószínűségszámítás és Statisztika

7. előadás

Arató Miklós

2020.03.31.

Tétel [Centrális határeloszlástétel független, azonos eloszlású változókra]: Legyenek
$\xi_1, \xi_2, \ldots$ független, azonos eloszlásúak, $m := E\xi_1$ és
$0 < \sigma^2 = D^2\xi_i < \infty$. Ekkor minden $x \in \mathbb{R}$-re

$$P\left(\frac{\sum\limits_{i=1}^{n} \xi_i - n \cdot m}{\sigma \cdot \sqrt{n}} < x\right) \to \Phi(x) = \int\limits_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}} \, dt$$

Példa 1.

Egy országban meg szeretnék becsülni a koronavírussal fertőzöttek arányát úgy, hogy legalább $0,95$ valószínűséggel legfeljebb 1%-ot tévedjünk.

Példa 1.

Egy országban meg szeretnék becsülni a koronavírussal fertőzöttek arányát úgy, hogy legalább $0,95$ valószínűséggel legfeljebb 1%-ot tévedjünk. Jelölje N az ország lakosai, M a koronavírusosak szavazók, n pedig a megvizsgáltak számát, ekkor $p := \frac{M}{N}$-et akarjuk jól közelíteni. Legyen továbbá X_i értéke 1, ha az i-edik megvizsgált koronavírusos és 0 különben. Ekkor a

$$P\left(\left|\frac{x_1 + \ldots + x_n}{n} - \frac{M}{N}\right| \leq 0,01\right) \geq 0,95$$

egyenlőtlenségnek kell teljesülnie.

Példa 1. (folyt.)

A centrális határeloszlástétel szerint

$$P\left(\frac{-\sqrt{n}\cdot 0,01}{\sqrt{p(1-p)}} \le \frac{\sum_{i=1}^{n} x_i - np}{\sqrt{n}\sqrt{p(1-p)}} \le \frac{\sqrt{n}\cdot 0,01}{\sqrt{p(1-p)}}\right) \sim$$

$$\sim \Phi\left(\frac{\sqrt{n}\cdot 0,01}{\sqrt{p(1-p)}}\right) - \Phi\left(\frac{-\sqrt{n}\cdot 0,01}{\sqrt{p(1-p)}}\right) = 2\cdot\Phi\left(\frac{\sqrt{n}\cdot 0,01}{\sqrt{p(1-p)}}\right) - 1 \ge 0,95$$

azaz $\Phi\left(\frac{\sqrt{n}\cdot 0,01}{\sqrt{p(1-p)}}\right) \ge 0,975 = \Phi(1,96)$, tehát legyen $\frac{\sqrt{n}\cdot 0,01}{\sqrt{p(1-p)}} \ge 1,96$. Ezzel $\sqrt{n} \ge 1,96 \cdot \frac{1}{0,01} \cdot \sqrt{p(1-p)}$, ha [1] $n \ge 10000$, tehát ha normális közelítéssel dolgozunk, kb. 10 000 embert kell megkérdezni.

[1] mivel a $\sqrt{p(1-p)}$ nem lehet nagyobb $0,5$-nél, így a jobb oldal felülről becsülhető 98-cal

Példa 2.

Mihez tart a következő: $e^{-n}\sum_{k=0}^{n-1}\frac{n^k}{k!}\xrightarrow{n\to\infty}?$

Ha végtelenig összegeznénk, nyilván 1 lenne az összeg, de most $n-1$-ig megyünk!

Példa 2. (folyt.)

Legyenek $\eta_i \sim$ 1-Poisson függetlenek, ezekre teljesülnek a centrális határeloszlástétel feltételei, azaz felírható:

$P\left(\frac{\sum_{i=1}^{n} \eta_i - n \cdot 1}{1 \cdot \sqrt{n}} < x\right) \to \Phi(x)$, ahol $\sum_{i=1}^{n} \eta_i \sim n$-Poisson, így

$P\left(\sum_{i=1}^{n} \eta_i < n\right) = e^{-n} \cdot \sum_{k=0}^{n-1} \frac{n^k}{k!}$.

Ekkor speciálisan $x = 0$-ra $P\left(\frac{\sum_{i=1}^{n} \eta_i - n}{\sqrt{n}} < 0\right) \to \Phi(0) = \frac{1}{2}$.

Tehát a keresett összeg $\frac{1}{2}$-hez tart.