

Magistrsko dela

Jaša Štefan

13. april 2020

1 CLI

Izrek 1. *centralni limitni izrek - osnovna (klasična) verzija*

Naj bo $\{X_1, X_2, \dots, X_n\}$ zaporedje n neodvisnih in enako porazdeljenih slučajnih spremenljivk z matematičnim upanjem μ in varianco σ . Naj bo slučajna spremenljivka Z definirana kot $Z := \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$. Potem je v limiti, ko $n \rightarrow \infty$, slučajna spremenljivka Z porazdeljena standardno normalno.

Izrek nam pove, da tudi če delamo z drugače (nenormalno) porazdeljenimi slučajnimi spremenljivkami, njihova standardizirana vsota vseeno konvergira k standardno normalni porazdelitvi. Zaradi tega lahko za statistično modeliranje te vsote uporabimo orodja in metode, katere uporabljamo že pri standardno normalnih porazdelitvah.

2 Kolmogorov - Smirnova razdalja

3 Berry - Esseenov izrek

Berry - Esseenov izrek oziroma neenakost nam pove, kako hitro porazdelitev standardiziranega povprečja konvergira proti normalni porazdelitvi, s tem ko omeji največjo napako aproksimacije med obema porazdelitvama. Točnost aproksimacije se meri z Kolmogorov-Smirnovo razdaljo.

V primeru, da gre za neodvisne vzorce (slučajne spremenljivke), je hitrost konvergence $n^{-\frac{1}{2}}$. Konstanto ocenimo s pomočjo koeficienta simetrije. Trenutno je zgornja meja za konstanto 0,4748, spodnja meja pa 0,40973.

Izrek je zelo močan tudi zato, ker potrebujemo samo prve tri centralne momente.

Izrek 2. Berry - Esseenov izrek Naj bodo X_1, X_2, \dots, X_n neodvisne in enako porazdeljene slučajne spremenljivke z $E[X_1] = 0$, $E[X_1^2] = \sigma^2$ in $E[X_1^3] = \rho$, pri čemer $\sigma > 0$ in $\rho < \infty$. Definirajmo slučajno spremenljivko Y_n , ki je vzorčno povprečje, $Y_n = \frac{X_1 + X_2 + \dots + X_n}{n}$. Nadalje definirajmo komulativno porazdelitveno funkcijo F_n slučajne spremenljivke $\frac{Y_n}{\sigma\sqrt{n}}$, s Φ pa označimo komulativno porazdelitveno funkcijo standardno normalne porazdelitve.

Potem obstaja konstanta $C, C > 0$, da velja

$$|F_n(x) - \Phi(x)| \leq \frac{C\rho}{\sigma^3\sqrt{n}}.$$

Primer 1. RULETA Igramo igro, kjer so možne vrednosti $0, 1, \dots, 36$ in je verjetnost vsake izmed njih enaka. Vsakič napovemo 1 cifro. V primeru, da zadenemo, dobimo 35 enot, sicer 1 enoto izgubimo. Igro ponovimo tisočkrat, rezultati so med seboj neodvisni.

Naj X_1, X_2, \dots, X_3 označujejo dobiček v vsaki igri.

Izračunamo $E[X_1] = -\frac{1}{37}$, $\text{var}(X_1) = 34,080$ in $\text{skew}(X_1) = 1162,366$.

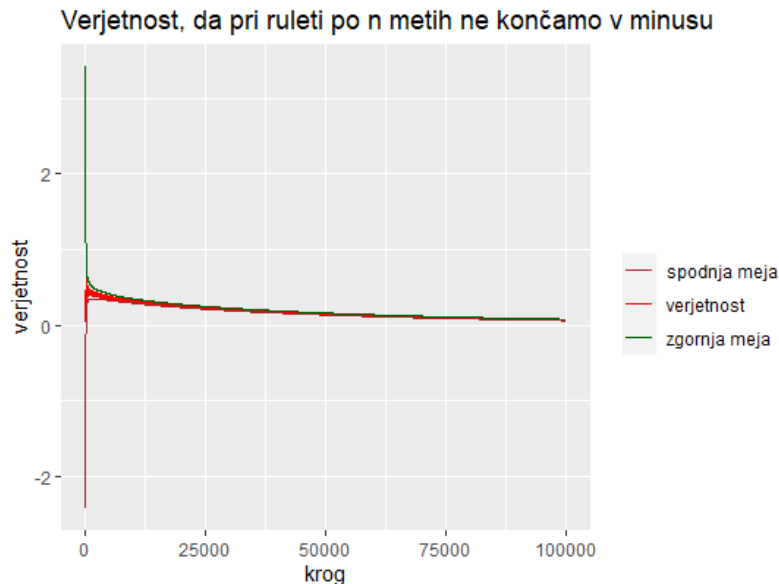
Definiramo $Y_n = \frac{X_1 + X_2 + \dots + X_{1000} - n\mu}{\sigma\sqrt{n}}$, F_n pa naj bo definirana kot zgoraj.

Sledi

$$|F_n(x) - \Phi(x)| \leq \frac{C \cdot 1162,366}{\sigma^3\sqrt{n}}.$$

Za $C = 0.5$ dobimo rezultat

$$|F_n(x) - \Phi(x)| \leq \frac{2.921}{\sqrt{n}}.$$



primerjava vzorca z normalno porazdelitvijo

