A május 9-i gyakorlat témája
Rövid összefoglaló

Megbeszéltünk néhány feladatot, mely arra adott példát, hogy bizonyos valószínűségeket hogyan lehet jól megbecsülni a centrális határeloszlástétel segítségével. Előtte felidéztük, hogy a $\Phi(\cdot)$ normális eloszlásfüggvény teljesíti a $\Phi(-x) = 1-\Phi(x)$ azonosságot tetszőleges $-\infty < x < \infty$ számra, és ez lehetővé teszi, hogy elég legyen a normális eloszlásfüggvényt csak nem negatív számokra tabulálni a táblázatokban.

Az említett azonosság abból következik, hogy a normális eloszlásfüggvény $\varphi(\cdot)$ eloszlásfüggvénye páros, ezért $\Phi(-x) = \int_{-\infty}^{-x} \varphi(u)\,du = \int_x^\infty \varphi(u)\,du = 1-\Phi(x)$.

Vizsgáljuk a következő feladat megoldását:

1. *Tekintsünk egy szabályos pénzdarab 10 000 egymás utáni (független) feldobásából származó fej-írás sorozatot. Adjunk becslést a Chebishev egyenlőtlenség segítségével annak valószínűségére, hogy a fej-dobások számának eltérése a várt 5000 számtól legalább 100-zal, illetve legalább 200-zal eltér! Milyen becslést ad ezekre a valószínűségekre a centrális határeloszlástétel?*

Megoldás: Vezessük be a következő ξ_j, $\le j \le 10\,000$ valószínűségi változókat, melyekre $\xi_j = 1$, ha a j-ik dobás eredménye fej, $\xi_j = 0$, ha a j-ik dobás eredménye írás. Ekkor ξ_j, $1 \le j \le 10\,000$ független valószínűségi változők, $E\xi_j = \frac12$, $\mathrm{Var}\,\xi_j = \frac14$, és a $P\left(\left|\sum\limits_{j=1}^{10000}(\xi_j - E\xi_j)\right| > 100\right)$ és $P\left(\left|\sum\limits_{j=1}^{10000}(\xi_j - E\xi_j)\right| > 200\right)$ valószínűségekre kell becslést adnunk. A Chebishev egyenlőtlenség az első valószínűségre a $\dfrac{10000\cdot\frac14}{100^2} = \dfrac14$, a második valószínűségre pedig a $\dfrac{10000\cdot\frac14}{200^2} = \dfrac{1}{16}$ felső becslést adja.

A centrális határeloszlástétel szerint $P\left(\dfrac{\sum\limits_{j=1}^{10000}(\xi_j - E\xi_j)}{\sqrt{10000\cdot\frac14}} > u\right) \sim 1-\Phi(u)$. Innen kapjuk, hogy az első valószínűség kiszámításához az $u = \pm\frac{100}{\frac12\cdot 100} = \pm 2$ értékeket kell tekinteni, és a vizsgált valószínűség közelítőleg $(1-\Phi(2)) + \Phi(-2) = 2(1-\Phi(2)) \sim 2(1-0.97720) = 0.0456$. A második valószínűség hasonlóan körülbelül $2(1-\Phi(4)) \sim 0$, (az első 4 tizedesjegy 0).

2. Legyen birtokunkban 100 lámpa, melyek mindegyike egymástól független időtartamig működik, élettartamuk pedig exponenciális eloszlású $\lambda = \frac{1}{100}$ paraméterrel. (A lámpák élettartamának exponenciális eloszlása természetes feltételezés.) Egy termet bevilágítunk ezen lámpák valamelyikével, majd amikor az kiégett új lámpát használunk fel. Adjunk jó becslést arra, hogy a lámpák összélettartama legalább 1150 óra.

Jelölje ξ_j a j-ik lámpa éléttartamát, $1 \le j \le 100$. Ekkor a $P(\xi_1 + \cdots + \xi_{100} >$

1150) valószínűségre kell jó becslést adnunk, ahol az összegben független exponenciális eloszlású valószínűségi változók szerepelnek $\lambda = \frac{1}{10}$ paraméterrel. Vezessük be az $\eta = \xi_1 + \cdots + \xi_{100}$ jelölést. Ki tudjuk-e számítani a $P(\eta > x)$ valószínűségeket tetszőleges x számra pontosan? A válasz erre a kérdésre az, hogy igen, de a jelen feladat megoldásához nem érdemes ezt az útat választani.

Ugyanis ki tudjuk számítani m független λ paraméterű exponenciális eloszlású valószínűségi változő összegének a sűrűségfüggvényét. Ez ugyanazt jelenti mint m darab exponenciális sűrűségfüggvény konvolucióját kiszámítani, és ezt a feladatot az április 25-i gyakorlat 3. feladatában tárgyaltuk. Eszerint esetünkben a keresett sűrűségfüggvény $\bar{f}(u) = \dfrac{\lambda^m u^{m-1}}{(m-1)!} e^{-\lambda u}$ $\lambda = \dfrac{1}{10}$ és $m = 100$ választással. Továbbá $P(\eta > x) = \int_x^\infty \bar{f}(u)\,du$. Ezt az integrált szukcessziv parciális integrálással ki tudjuk számolni, és kapunk egy száztagú összeget. Ez egy elvi lehetőséget ad a minket érdeklő valószínűség kiszámítására $x = 1150$ választással. Ez azonban sok számolást igényel, és félő hogy a kerekítési hibák miatt a kapott eredmény félrevezető lesz még hibátlan számolás esetén is. Ennél egyszerűbb és hasznosabb a centrális határeloszlástétel alkalmazása.

Kiszámoltuk (április 25-i gyakorlat 2. feladat), hogy jelen esetben $E\eta = mE\xi_1 = \dfrac{m}{\lambda} = 10$, $\operatorname{Var}\eta = \dfrac{m}{\lambda^2} = 10000$ ($m = 100$ és $\lambda = \dfrac{1}{10}$ választással. Ezért a centrális határeloszlástétel szerint $\dfrac{\eta - E\eta}{\sqrt{\operatorname{Var}\eta}} = \dfrac{\eta - 1000}{100}$ jó közelítéssel standard normális eloszlású valósz1nűségi változó, és $P(\xi_1 + \cdots + \xi_{100} > 1150) = \dfrac{\eta - E\eta}{\sqrt{\operatorname{Var}\eta}} > 1.5 \sim 1 - \Phi(1.5)$.

3. Vegyünk egy olyan pénzdarabot, mely $\frac{2}{3}$ valószínűséggel esik a fej és $\frac{1}{3}$ valószínűséggel az írás oldalra. Ezt a pénzdarabot annyiszor dobjuk fel, ameddig megjelenik 1200 fej dobás. Mi annak a valószínűsége, hogy az elvégzett dobások száma 1680 1430 küzé esik? Adjunk erre a valószínűségre jó közelítő becslést.

Az elvégzett dobások száma negatív binomiális eloszlású valószínűségi változó $n = 1200$ $p = \frac{1}{3}$ paraméterekkel, és egy ilyen valószínűségi változónak kiszámoltuk a pontos eloszlását, azaz azt, hogy milyen értéket milyen valószínűséggel vesz fel. (Lásd március 28-i gyakorlat 6. feladat.) Elvileg, ez lehetőséget ad a kívánt valószínűség kiszámítására. Ennél hasznosabb becslést tudunk kapni a következő érvelés segítségével, mely a kívánt valószínűséget jó pontossággal kiszámítja a centrális határeloszlástétel segítségével.

Jelölje ξ_j, $2 \le j \le 1200$ a $j-1$-ik és j-ik fejdobás közötti dobások számát (a j-ik fejdobást beleszámítjuk a $j-1$-iket viszont nem számítjuk bele e dobások közé), és legyen ξ_1 az első fejdobásig (ezt is beleszámítva) elvégzett dobások száma. Ekkor a ξ_j valószínűségi változók függetlenek, negatív binomiális eloszlásúak $n = 1$, $p = \frac{1}{3}$ paraméterekkel, és minket a $P(1680 < \xi_1 + \cdots + \xi_{1200} < 1830)$ valószínűség érdekel. Megmutattuk (március 28-i gyakorlat 6. feladat), hogy $E\xi_j = \dfrac{1}{1-p} = 3$, $\operatorname{Var}\xi_j = \dfrac{p}{(1-p)^2} = 6$. Ezért a centrális határeloszlástétel alapján $\eta = \xi_1 + \cdots + \xi_{1200}$

jelöléssel minket a $P\left(-\sqrt{2} < \frac{\eta - 1200 E\xi_1}{\sqrt{1200 \operatorname{Var} \xi_1}} < \frac{\sqrt{2}}{4}\right)$ valószínűség érdekel. A centrális határeloszlástétel alapján $P\left(-\sqrt{2} < \frac{\eta - 1200 E\xi_1}{\sqrt{1200 \operatorname{Var} \xi_1}} < \frac{\sqrt{2}}{4}\right) \sim \Phi\left(\frac{\sqrt{2}}{4}\right) + \Phi(\sqrt{2}) - 1$.

Még volt néhány egyéb probléma tárgyalása is, melyeket csak később írok le. A le nem írt anyag nem fog szerepelni a május 16-án írandó dolgozatban.