**Valószínűségi becslések, alapvető technikák**

Legyen adva $\xi_1$, $\xi_2$, … $\xi_n$ független egyforma eloszlású független valószínűségi változók sorozata. Kiváncsiak vagyunk az összeg eloszlásának aszimptotikus viselkedésére, arra hogy a „szép esetekben” milyen pontosságú becslést adhatunk erre. Át kívánjuk tekinteni azokat az alapvető technikákat, melyek segítségével erre a kérdésre választ tudunk adni, és melyik módszer mire alkalmas. Feltesszük, hogy a $\xi$ valószínűségi változók eloszlása szép, $E\xi_1 = 0$, $E\xi_1^2 = 1$, az összes momentum létezik, és a $\xi_1$ valószínűségi változó $F(x)$ eloszlásfüggvényének létezik szép $f(x)$ sűrűségfüggvénye.

Foglalkozzunk először az $S_n = \sum\limits_{k=1}^{n} \xi_k$, $f_n(x) = f^{*(n)}(x)$, sűrűségfüggvényének az aszimptotkus viselkedésével, ahol $f^{*(n)}$ az $f$ függvény $n$-szeres konvolúcióját jelenti önmagával. Ezt ugyanis megadhatjuk a Fourier analízis néhány alapvető eredményének a segítségével. Az eloszlásfüggvény aszimptotikus viselkedését szintén megkaphatjuk e technika segítségével. Némi simítással, amennyiben az szükséges, elérhetjük, hogy olyan eloszlás konvolúcióját elég vizsgálni, melynek van sűrűségfüggvénye, és ha az összeg sűrűségfüggvényére van jó aszimptotika, akkor azt kiintegrálva, az eloszlásfüggvényre is jó aszimptotikát kapunk.

Idézzük fel a következő eredményeket: Ha $f(x)$ tetszőleges integrálható függvény, amelyiknek a Fourier transzformáltja

$$\tilde{f}(t) = \int_{-\infty}^{\infty} e^{itx} f(x)\, dx$$

is integrálható (ez $f(x)$-re simasági feltétel), akkor az inverz transzformáció a következő:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \tilde{f}(t)\, dt\ .$$

*Feladat:* Értsük meg és bizonyítsuk be a fenti formulát. Használjuk fel, hogy egy sima periódikus függvény egyenlő a Fourier sorával. Ezt a tényt kihasználva egy $[-T, T]$ intervallumban, majd alkalmazva a $T \to \infty$ határátmenetet bizonyítsuk be az állítást sima $f$ függvényekre. Simítás segítségével szabaduljunk meg a simasági feltételtől. Érdemes a simítást konvolúció segítségével végezni.

Vezessük be a $*$ konvolúció operátort: $f * g(x) = \int_{-\infty}^{\infty} f(y)g(x-y)\, dy$. Ekkor $\tilde{f}_n(t) = \tilde{f}^n(t)$, azaz függvények konvolúciójának a Fourier transzformáltja egyenlő a Fourier transzformáltak szorzatával, $\tilde{f}(t) \to 0$, ha $|t| \to \infty$, sőt $\tilde{f}(t) = O(|t|^{-k})$ vagy $O(e^{-\alpha|t|})$, ha az $f(x)$ sűrűségfüggvény elég sima, és $\sup\limits_{|t|\ge\varepsilon} |\tilde{f}(t)| < 1$ minden $\varepsilon > 0$-ra. Továbbá, $\left.\dfrac{d^k}{dt^k}\tilde{f}(t)\right|_{t=0} = i^k E\xi^k$. Miért igazak ezek az állítások?

Ezért

$$f^{*(n)}(x) = \frac{1}{2\pi} \int e^{-itx} \tilde{f}^n(t)\, dt$$

$$\frac{1}{\sqrt{n}} f^{*(n)}\left(\frac{x}{\sqrt{n}}\right) = \frac{1}{2\pi}\int e^{-itx}\tilde{f}^n\left(\frac{t}{\sqrt{n}}\right)\,dt$$

Innen kiolvasható a centrális határeloszlástétel lokális alakja, (azaz a sűrűségfüggvény konvergenciája a normális sűrűségfüggvényhez), és meg lehet adni a konvergencia sebességét. Valóban,

$$\begin{aligned}\tilde{f}^n\left(\frac{t}{\sqrt{n}}\right) = \left(1-\frac{t^2}{2n}+O\left(\frac{t^3}{n^{3/2}}\right)\right)^n &= e^{-n(t^2n^{-1}+O(n^{-3/2}t^3))}\\ &= e^{-t^2/2}\left(1+O\left(\frac{t^3}{\sqrt{n}}\right)\right).\end{aligned}$$

Ezért felhasználva a Fourier transzformált gyors eltünését a végtelenben kapjuk, hogy

$$\frac{1}{\sqrt{n}} f^{*(n)}\left(\frac{x}{\sqrt{n}}\right) = \frac{1}{2\pi}\int e^{-itx}e^{-t^2/2}\,dt + O(n^{-1/2}) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}+O(n^{-1/2}).$$

1.) Dolgozzuk ki e bizonyítást alkalmas feltételek teljesülése esetén.

Lehet a pontosabb $\tilde{f}(t) = 1-\frac{t^2}{2}+\frac{i^3m_3}{6}t^3+\cdots+\frac{i^km_k}{k!}t^k+O(t^{k+1})$ Taylor sorfejtést alkalmazni. Milyen pontosabb aszimptotikát adhatunk ennek segítségével? Bizonyítsuk be a következő állítást:

2.)

$$\int e^{itx}t^ke^{-t^2/2}\,dt = \frac{i^k}{k!}\frac{1}{\sqrt{2\pi}}H_k(x)e^{-x^2/2},$$

ahol $H_k(x)$ a $k$-ik Hermite polinom 1 főegyütthatóval.

3.) Bizonyítsuk be, (alkalmas feltevések mellett), hogy

$$\begin{aligned}\frac{1}{\sqrt{n}} f^{*(n)}\left(\frac{x}{\sqrt{n}}\right) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}\bigg(1+d_3\frac{H_3(x)}{n^{1/2}}+d_4\frac{H_4(x)}{n}+\cdots\\ +d_k\frac{H_k(x)}{n^{(k-2)/2}}\bigg)+O(n^{-(k-1)/2}),\end{aligned}$$

ahol $H_k(x)$ a $k$-ik Hermite polinom, és a $d_j$ konstans a $\xi$ valószínűségi változó első $j$ momentumától függ minden $3\le j\le k$-ra.

4.) Tekintsük az $f(x)=e^{-x}$, ha $x\ge 0$, $f(x)=0$ ha $x<0$ (az exponenciális eloszláshoz tartozó) sűrűségfüggvényt. Írjuk fel erre a (lokális) centrális határeloszlástételt, és ennek segítségével bizonyítsuk be a Stirling formulát, (ami $n!$ aszimptotikus viselkedésének a leírása.) A (lokális) centrális határeloszlástételre adott sorfejtés segítségével adjunk sorfejtést a Stirling formulára is.

Ha a sűrűségfüggvény helyett az eloszlásfüggvény aszimptotikus viselkedését akarjuk vizsgálni, akkor némi simításra van szükség. Ehhez hasznos a következő feladat megoldása.

5.) Létezik olyan $\mu$ valószínűségi mérték, amelyiknek a $\varphi(t) = \int_{-\infty}^{\infty} e^{itx}\,\mu(dx)$ Fourier transzformáltja (karakterisztikus függvénye) $|t| > 1$ re eltünik, azaz $\varphi(t) = 0$ ha $|t| \geq 1$.

*Segítség:* (Egy lehetséges konstrukció.) Ha $f(x)$ szép, korlátos tartójú sűrűségfüggvény, akkor $f * f^-(x)$, $f^-(x) = f(-x)$, is az. Továbbá ennek Fourier transzformáltja $\widetilde{f * f^-}(x)$ nem negatív valós függvény. Ennek a ténynek és az inverz Fourier transzformáltra adott formulának a segítségével mutassuk meg, hogy $\widetilde{f * f^-}(x)$ alkalmas lineáris transzformációja választható mint egy kívánt tulajdonságú $\mu$ sűrűségfüggvénye.

6.) Legyen $F$ és $G$ két eloszlásfüggvény (vagy általánosabban $G$ olyan korlátos változású függvény, melyre $\lim\limits_{t\to-\infty} G(t) = 0$, $\lim\limits_{t\to\infty} G(t) = 1$), $\omega(x)$ az 5. feladat feltételeit kielégítő $\mu$ mérték sűrűségfüggvénye, $\omega_T(x) = \frac{1}{T}\omega\left(\frac{x}{T}\right)$, $F_T(x) = F * \omega_T(x)$, $G_T(x) = G * \omega_T(x)$. Bizonyítsuk be, hogy

$$F_T(x) - G_T(x) = \frac{1}{2\pi}\int_{-T}^{T} e^{-itx}\frac{\tilde{F}(t) - \tilde{G}(t)}{-it}\tilde{\omega}_T(t)\,dt\ ,$$

$$\Delta_T(x) = |F_T(x) - G_T(x)| \leq \frac{1}{2\pi}\int_{-T}^{T}\frac{|\tilde{F}(t) - \tilde{G}(t)|}{|t|}\,dt$$

ahol ˜ Fourier transzformációt jelöl.

Kérdések: A minket érdeklő esetekben ($F$ független valószínűségi változok összegének sztandardizáltjának az eloszlásfüggvénye, $G(x)$ a normális eloszlásfüggvény vagy a harmadik feladat jobboldalán szereplő függvény integrálja) miért nem zavar az, hogy a becslésben $t$-vel osztunk? Miért jobb ez a becslés annál, amit az $F_T - G_T$ deriváltjának a kiintegrálásából a fent ismertetett becslések segítségével kapnánk?

7.) Ha jó becslésünk van $\Delta_T(x) = |F_T(x) - G_T(x)|$-re viszonylag nagy $T$-vel, akkor hogyan tudunk jó becslést kapni a $\Delta(x) = |F(x) - G(x)|$ kifejezésre?

(A fent ismertett becslések a sűrűség és eloszlásfüggvényre csak az $|x| \leq \text{const.}\sqrt{\log n}$ intervallumban hasznosak. Miért? Hogyan lehet jó aszimptotikát adni, ha $x \gg \sqrt{\log n}$? E probléma vizsgálatában hasznos a nyeregpont módszer.)