**A Valószínűségszámítás II. előadássorozat negyedik témája.**

## A CENTRÁLIS HATÁRELOSZLÁSTÉTEL

A valószínűségszámítás legfontosabb eredménye a centrális határeloszlástétel. Ez azt mondja ki, hogy független valószínűségi változók normalizált részletösszegeinek az eloszlása nagyon általános feltételek teljesülése esetén közelítőleg a standard normális eloszlás. Megfogalmazom, majd később bebizonyítom a centrális határeloszlástétel legáltalánosabb alakját. Ezután ismertetek néhány fontos eredményt, amelyeket ez a tétel speciális esetként tartalmaz. Mutatok néhány példát, amelyek magyarázatot adnak a tételben megfogalmazott feltételek szerepére. Megfogalmazom, és a kiegészítésben bebizonyítom a centrális határeloszlástétel megfordítását is. Ez az eredmény azt állítja, hogy a tételben megfogalmazott feltételek nemcsak elégséges, hanem egyben szükséges feltételei is a centrális határeloszlástételnek.

Röviden tárgyalom az úgynevezett lokális centrális határeloszlástételt is. Ez azt mondja ki, hogy alkalmas feltételek teljesülése esetén független valószínűségi változók normalizált összegeinek nemcsak az eloszlásfüggvényük van közel a standard normális eloszlásfüggvényhez, hanem a sűrűségfüggvényük is közel van a standard normális sűrűségfüggvényhez. Ennek az eredménynek a bizonyításában hasonló gondolatok jelennek meg, mint amilyenekkel a Stirling formula korábban ismertetett bizonyításában találkoztunk.

Az eredmények ismertetése előtt bevezetek néhány fogalmat. A centrális határeloszlástétel általános alakja szériasorozatok részletösszegeinek eloszlásáról szól. Ezért ismertetem a szériasorozatok definicióját.

**Szériasorozatok definiciója.** *A*

$$\begin{matrix} \xi_{1,1},\dots,\xi_{1,n_1} \\ \vdots \qquad \vdots \\ \xi_{k,1},\dots,\xi_{k,n_k} \\ \vdots \qquad \vdots \end{matrix}$$

*valószínűségi változók rendszere, $k\to\infty$, szériasorozat, ha az egy sorban levő $\xi_{k,1}$, … , $\xi_{k,n_k}$ valószínűségi változók függetlenek. (A különböző sorokban levő valószínűségi változók kapcsolatáról nem tételezünk fel semmit.)*

A centrális határeloszlástétel szériasorozatokra azt állítja, hogy egy $\xi_{k,j}$, $k=1$, $2$, … , $1\le j\le n_k$, szériasorozat $S_k=\sum\limits_{j=1}^{n_k}\xi_{k,j}$, $k=1,2,\dots$, sorösszegei alkalmas feltételek teljesülése esetén eloszlásban konvergálnak a standard normális eloszláshoz. Természetes megkövetelni, hogy nagy $k$ indexre a $\xi_{k,j}$, $1\le j\le n_k$, sornak ne legyen olyan eleme, amely ugyanolyan nagyságrendű, mint a többi tag összege együttvéve. Ilyen követelményt fogalmaz meg az egyenletes kicsiség alább bevezetett feltétele.

**Egyenletes kicsiség feltétele szériasorozatokra.** *Legyen* $\xi_{k,j}$, $k = 1, 2, \ldots$, $1 \le j \le n_k$, *olyan szériasorozat, amelyre* $E\xi_{k,j} = 0$, $E\xi_{k,j}^2 < \infty$, $k = 1, 2, \ldots$, $1 \le j \le n_k$, *és teljesül a* $\lim\limits_{k\to\infty} \sum\limits_{j=1}^{n_k} E\xi_{k,j}^2 = 1$ *reláció. Egy ilyen szériasorozat teljesíti az egyenletes kicsiség feltételét, ha*

$$\lim_{k\to\infty} \left( \sup_{1\le j\le n_k} E\xi_{k,j}^2 \right) = 0.$$

A centrális határeloszlástétel általános (szériasorozatok sorösszegeiről szóló) alakjában fontos szerepet játszik az alábbi Lindeberg feltétel.

**Lindeberg feltétel definiciója szériasorozatokra:** *Legyen* $\xi_{k,j}$, $k = 1, 2, \ldots$, $1 \le j \le n_k$, *olyan szériasorozat, amelyre* $E\xi_{k,j} = 0$, $E\xi_{k,j}^2 < \infty$, $k = 1, 2, \ldots$, $1 \le j \le n_k$, *és* $\lim\limits_{k\to\infty} \sum\limits_{j=1}^{n_k} E\xi_{k,j}^2 = 1$. *Ez a szériasorozat akkor és csak akkor teljesíti a Lindeberg feltételt, ha tetszőleges* $\varepsilon > 0$ *számra*

$$\lim_{k\to\infty} \sum_{j=1}^{n_k} E\xi_{k,j}^2 I\left(\{|\xi_{k,j}| > \varepsilon\}\right) = 0,$$

*ahol* $I(A)$ *egy* $A$ *halmaz indikátor függvénye.*

Megfogalmazom a centrális határeloszlástétel általános alakját szériasorozatok sorösszegeire.

**A centrális határeloszlástétel szériasorozatok sorösszegeire.** *Legyen* $\xi_{k,j}$, $k = 1, 2, \ldots$, $1 \le j \le n_k$, *olyan szériasorozat, amelynek tagjaira* $E\xi_{k,j} = 0$, $E\xi_{k,j}^2 < \infty$, $k = 1, 2, \ldots$, $1 \le j \le n_k$, $\lim\limits_{k\to\infty} \sum\limits_{j=1}^{n_k} E\xi_{k,j}^2 = 1$, *és teljesítse ez a szériasorozat a Lindeberg feltételt. Ekkor*

*a.) A szériasorozat teljesíti a* $\lim\limits_{k\to\infty} \left( \sup\limits_{1\le j\le n_k} E\xi_{k,j}^2 \right) = 0$ *egyenletes kicsiség feltételét.*

*b.) Az* $S_k = \sum\limits_{j=1}^{n_k} \xi_{k,j}$, $1 \le k < \infty$, *véletlen összegek eloszlásban konvergálnak a standard, azaz nulla várható értékű és 1 szórásnégyzetű normális eloszláshoz, ha* $k \to \infty$.

Igaz a szériasorozatok sorösszegeiről szóló centrális határeloszlástétel következő megfordítása is.

**A szériasorozatok sorösszegeiről szóló centrális határeloszlástétel megfordítása.** *Legyen* $\xi_{k,j}$, $k = 1, 2, \ldots$, $1 \le j \le n_k$, *olyan szériasorozat, amelyre* $E\xi_{k,j} = 0$, $E\xi_{k,j}^2 < \infty$, $k = 1, 2, \ldots$, $1 \le j \le n_k$, $\lim\limits_{k\to\infty} \sum\limits_{j=1}^{n_k} E\xi_{k,j}^2 = 1$, *és teljesíti a szériasorozatokra* *megfogalmazott egyenletes kicsiségi feltételt is. Ha ez a szériasorozat teljesíti a centrális határeloszlástételt, azaz az* $S_k = \sum\limits_{j=1}^{n_k} \xi_{k,j}$, $1 \le k < \infty$, *véletlen összegek eloszlásban konvergálnak az egy szórásnégyzetű és nulla várható értékű normális eloszláshoz* $k \to \infty$ *esetén, akkor a* $\xi_{k,j}$, $1 \le k < \infty$, $1 \le j \le n_k$, *szériasorozat teljesíti a Lindeberg feltételt.*

*Megjegyzés:* A fenti eredményben a centrális határeloszlástétel teljesülésének a feltétele nemcsak azt jelenti, hogy az $S_k$ sorösszegek eloszlásban konvergálnak egy normális eloszláshoz, hanem azt is, hogy ennek a normális határeloszlásnak a '*helyes*' 1 szórásnégyzete van. Fogunk példát látni olyan az egyenletes kicsiség feltételét teljesítő szériasorozatokra, amelyek sorösszegei eloszlásban egy 'rossz' (túl kicsi szórásnégyzetű) normális eloszláshoz konvergálnak és nem teljesítik a Lindeberg feltételt. Azt a feltételt, hogy a határeloszlás várható értéke legyen nulla csak kényelmi okokból tettem fel. A tétel bizonyításából látható, hogy ez a feltétel elhagyható. Ha a szériasorozat teljesíti az egyenletes kicsiség feltételét és a sorösszegek egy egy szórásnégyzetű normális eloszlású valószínűségi változóhoz konvergálnak, akkor ez csak úgy lehetséges, hogy a határeloszlás a (nulla várható értékű) standard normális eloszlás.

A fenti tételek kielégítő magyarázatot adnak arra a kérdésre, hogy szériasorozatok sorösszegei mikor teljesítik a centrális határeloszlástételt. Minket azonban további a centrális határeloszlás problémaköréhez tartozó kérdések is érdekelnek. Például szeretnénk azt is tudni, hogy független valószínűségi változók normalizált részletösszegei mikor teljesítik a centrális határeloszlástételt. Részletesebben megfogalmazva ez a kérdés a következő:

Tekintsük független valószínűségi változók egy $\xi_n$, $n = 1, 2, \dots$, sorozatát, amelyre $E\xi_n = 0$, és $\sigma_n^2 = E\xi_n^2 < \infty$, $n = 1, 2, \dots$. Jelölje $S_n = \sum\limits_{k=1}^{n} \xi_k$ e sorozat első $n$ tagjának az összegét, és legyen $s_n^2 = \operatorname{Var} S_n = \sum\limits_{k=1}^{n} \sigma_k^2$ ezen véletlen összeg szórásnégyzete. Tegyük fel, hogy az $s_n$ számsorozat teljesíti a $\lim\limits_{n\to\infty} s_n^2 = \infty$ relációt. Azt szeretnénk tudni, hogy az $\frac{S_n}{s_n}$ normalizált részletösszegek mikor konvergálnak eloszlásban a standard normális eloszláshoz.

Nem nehéz látni, hogy ez a probléma speciális esete annak a kérdésnek, hogy szériasorozatok sorösszegei mikor teljesítik a centrális határeloszlástételt. Valóban, tekintsük független valószínűségi változók egy $\xi_n$, $n = 1, 2, \dots$, sorozatát, amelyre $E\xi_n = 0$, és $\sigma_n^2 = E\xi_n^2 < \infty$, $n = 1, 2, \dots$. Jelölje $S_n = \sum\limits_{k=1}^{n} \xi_k$ a sorozat első $n$ tagjának az összegét, és legyen $s_n^2 = \operatorname{Var} S_n = \sum\limits_{k=1}^{n} \sigma_k^2$ ezen véletlen összeg szórásnégyzete. Definiáljuk ezen $\xi_1, \xi_2, \dots$ független valószínűségi változók, illetve a segítségükkel bevezetett $s_n$, $n = 1, \dots$ mennyiségek felhasználásával a következő $\xi_{k,j}$, $k = 1, 2, \dots$, $j = 1, \dots, n_k$, szériasorozatot:

$n_k = k$ és $\xi_{k,j} = \frac{\xi_j}{s_k}$, ha $1 \le j \le k$ minden $k = 1, 2, \dots$ számra.

Ezzel a választással a az $\frac{S_n}{s_n}$ normalizált részletösszegek akkor és csak akkor konvergálnak eloszlásban a standard normális eloszláshoz, ha a $\xi_{k,j} = \frac{\xi_j}{s_k}$, $k = 1, 2, \dots, 1 \le j \le k$, szériasorozat sorösszegei teljesítik a centrális határeloszlástételt. Ezenkívül ezt a szériasorozatot úgy definiáltuk, hogy az teljesíti a $\sum\limits_{j=1}^{n_k} E\xi_{k,j}^2 = 1$, $k = 1, 2, \dots,$ relációt. Ezért a szériasorozatok sorösszegeiről szóló centrális határeloszlástétel természetes következményeként független valószínűségi változók normalizált részletösszegeire érvényes centrális határeloszlástételt kaphatunk. Ennek megfogalmazása érdekében vezessük be a Lindeberg feltétel és az egyenletes kicsiség feltételének természetes átfogalmazását független valószínűségi változók részletösszegeire.

**Az egyenletes kicsiség feltétele független valószínűségi változók sorozataira:** *Legyen* $\xi_n$, $n = 1, 2, \dots,$ *független valószínűségi változók sorozata, amelyre* $E\xi_n = 0$, $\sigma_n^2 = E\xi_n^2 < \infty$, $n = 1, 2, \dots.$ *Vezessük be az* $s_n^2 = \sum\limits_{k=1}^{n} \sigma_k^2$, $n = 1, 2, \dots,$ *mennyiségeket. Azt mondjuk, hogy a* $\xi_n$, $n = 1, 2, \dots,$ *sorozat teljesíti az egyenletes kicsiség feltételét, ha* $\lim\limits_{n\to\infty} s_n^2 = \infty$, *és* $\lim\limits_{n\to\infty} \sup\limits_{1\le k\le n} \frac{\sigma_k^2}{s_n^2} = 0.$

**Lindeberg feltétel független valószínűségi változók sorozataira:** *Legyen* $\xi_n$, $n = 1, 2, \dots,$ *független valószínűségi változók sorozata, amelyre* $E\xi_n = 0$, $\sigma_n^2 = E\xi_n^2 < \infty$, $n = 1, 2, \dots,$ *és az* $s_n^2 = \sum\limits_{k=1}^{n} \sigma_k^2$, $n = 1, 2, \dots,$ *sorozat teljesíti a* $\lim\limits_{n\to\infty} s_n^2 = \infty$ *feltételt. Azt mondjuk, hogy a* $\xi_n$, $n = 1, 2, \dots,$ *sorozat teljesíti a Lindeberg feltételt, ha minden* $\varepsilon > 0$ *számra*

$$\lim_{n\to\infty} \frac{1}{s_n^2} \sum_{k=1}^{n} E\xi_k^2 I(\{|\xi_k| > \varepsilon s_n\}) = 0.$$

A következő centrális határeloszlástétel, illetve annak megfordítása egyszerű következménye a szériasorozatok részletösszegeiről kimondott centrális határeloszlástételnek és annak megfordításának.

**A centrális határeloszlástétel független valószínűségi változók normalizált részletösszegeire.** *Legyen* $\xi_n$, $n = 1, 2, \dots,$ *független valószínűségi változók sorozata, amelyre* $E\xi_n = 0$, $\sigma_n^2 = E\xi_n^2 < \infty$, $n = 1, 2, \dots,$ $\lim\limits_{n\to\infty} s_n^2 = \infty$, *ahol* $s_n^2 = \sum\limits_{k=1}^{n} \sigma_k^2$. *Teljesítse ez a* $\xi_n$, $n = 1, 2, \dots,$ *sorozat a független valószínűségi változók összegeire megfogalmazott Lindeberg feltételt. Ekkor ez a sorozat teljesíti a független valószínűségi változók sorozataira megfogalmazott egyenletes kicsiségi feltételt, és az* $\frac{S_n}{s_n} = \frac{\sum_{k=1}^{n} \xi_k}{s_n}$ *normalizált részletösszegek eloszlásban konvergálnak a standard normális eloszláshoz, ha* $n \to \infty$.

**Független valószínűségi változók normalizált részletösszegeiről szóló centrális határeloszlástétel megfordítása.** *Legyen* $\xi_n$, $n = 1, 2, \dots,$ *független valószínűségi* *változók sorozata, amelyre $E\xi_n = 0$, $\sigma_n^2 = E\xi_n^2 < \infty$, $n = 1, 2, \dots,$ $\lim\limits_{n\to\infty} s_n^2 = \infty$, ahol $s_n^2 = \sum\limits_{k=1}^{n} \sigma_k^2$. Teljesítse ez a $\xi_n$, $n = 1, 2, \dots,$ sorozat a független valószínűségi változók összegeire megfogalmazott egyenletes kicsiség feltételét. Ha az $\frac{S_n}{s_n} = \frac{\sum_{k=1}^{n} \xi_k}{s_n}$ normalizált részletösszegek eloszlásban konvergálnak a standard normális eloszláshoz $n \to \infty$ esetén, akkor a $\xi_n$, $n = 1, 2, \dots$ sorozat teljesíti a független valószínűségi változók sorozataira megfogalmazott Lindeberg feltételt.*

Szeretnénk látni, hogy a fent megfogalmazott általános, centrális határeloszlástétel alkalmazható a minket érdeklő esetekben. A fő probléma az, hogy mikor teljesül a centrális határeloszlás feltételeként megfogalmazott Lindeberg feltétel. Tekintsük először azt a fontos, speciális esetet, amikor független, egyforma eloszlású valószínűségi változók normalizált részletösszegeit vizsgáljuk.

**Centrális határeloszlástétel független, egyforma eloszlású valószínűségi változók normalizált részletösszegeire.** *Legyen $\xi_1, \xi_2, \dots,$ független, egyforma eloszlású valószínűségi változók sorozata, amelyre $E\xi_1^2 < \infty$. Vezessük be a $\sigma^2 = \mathrm{Var}\,\xi_1 = E\xi_1^2 - (E\xi_1)^2$ mennyiséget. Ekkor az $\frac{S_n - nE\xi_1}{\sqrt{n}\sigma} = \frac{\sum_{k=1}^{n} \xi_k - nE\xi_1}{\sqrt{n}\sigma}$ normalizált összegek eloszlásban konvergálnak a standard normális eloszláshoz $n \to \infty$ esetén.*

*Bizonyítás.* Azt fogom megmutatni, hogy ez az eredmény megkapható a független valószínűségi változók normalizált részletösszegeire vonatkozó centrális határeloszlástétel speciális eseteként. Feltehetjük, hogy a $E\xi_1 = 0$, $\xi_1^2 = 1$. Valóban, bevezetve a $\xi_k' = \frac{\xi_k - E\xi_k}{\sigma_1}$, $k = 1, 2, \dots,$ mennyiségeket, egy független egyforma, eloszlású, nulla várható értékű és 1 szórásnégyzetű valószínűségi változókból álló $\xi_1', \xi_2' \dots$ sorozatot definiáltunk, amelyre $\frac{S_n'}{\sqrt{n}} = \frac{\sum_{k=1}^{n} \xi_k'}{\sqrt{n}} = \frac{S_n}{\sqrt{n}\sigma}$. Ezért elegendő a centrális határeloszlástételt csak erre az új esetre belátni.

Ha $\xi_1, \xi_2, \dots,$ független, egyforma eloszlású valószínűségi változók sorozata, amelyekre $E\xi_1 = 0$, $E\xi_1^2 = 1$, akkor a centrális határeloszlástétel bizonyításához elég belátni, hogy ez a sorozat teljesíti a Lindeberg feltételt. Mivel ebben ez esetben $s_n^2 = \sum\limits_{k=1}^{n} E\xi_k^2 = n$, ez azt jelenti, hogy azt kell megmutatni, hogy

$$\frac{1}{s_n^2} \sum_{k=1}^{n} E\xi_k^2 I(|\xi_k| > \varepsilon s_n) = \frac{1}{n} \sum_{j=1}^{n} E\xi_j^2 I\left(|\xi_j| > \varepsilon \sqrt{nE\xi_j^2}\right) = E(\xi_1^2 I(|\xi_1| > \varepsilon\sqrt{n}) \to 0$$

minden $\varepsilon > 0$ számra, ha $n \to \infty$. Viszont az $E(\xi_1^2 I(|\xi_1| > \varepsilon\sqrt{n}) \to 0$, ha $n \to \infty$ reláció következik a Lebesgue tételből, mert $\xi_1^2 I(|\xi_1| > \varepsilon\sqrt{n}) \to 0$ 1 valószínűséggel, ha $n \to \infty$, és $\xi_1^2(\omega) I(|\xi_1(\omega)| > \varepsilon\sqrt{n}) \le \xi_1^2(\omega)$ minden $\omega \in \Omega$ pontban és $n = 1, 2, \dots$ indexre.

Megfogalmazok egy további eredményt, amely azt mondja ki, hogy a centrális határeloszlástétel érvényes alkalmas feltételek teljesülése esetén. Ez az eredmény arról szól, hogy a Lindeberg feltétel teljesül, ha bizonyos könnyebben ellenőrizhető feltételek teljesülnek.

**A centrális határeloszlástétel függgetlen valószínűségi változók normalizált részletösszegeire.** *Legyen $\xi_n$, $n=1,2,\dots$, független valószínűségi változók sorozata, amelyre $E\xi_n=0$, $\sigma_n^2=E\xi_n^2<\infty$, $n=1,2,\dots$, $\lim\limits_{n\to\infty} s_n^2=\infty$, ahol $s_n^2=\sum\limits_{k=1}^n \sigma_k^2$. Ez a sorozat teljesíti a Lindeberg feltételt, így a $\frac{\sum\limits_{k=1}^n \xi_k}{s_n}$ normalizált összegek, $n=1,2,\dots$, eloszlásban konvergálnak a standard normális eloszláshoz, ha a következő feltétel teljesül.*

*a) $E|\xi_k|^{2+\alpha}<\infty$, minden $k=1,2,\dots$ számra valamilyen $\alpha>0$ konstanssal, és $\lim\limits_{n\to\infty} \frac{\sum\limits_{k=1}^n E|\xi_k|^{2+\alpha}}{s_n^{2+\alpha}}=0$.*

*Speciálisan, ez a feltétel teljesül akkor, ha $E\xi_k^2\ge K$ valamilyen $K>0$ számmal minden $k=1,2,\dots$ indexre, és ezenkívül érvényes a $\lim\limits_{k\to\infty} k^{-\alpha/2}E|\xi_k|^{2+\alpha}=0$ reláció.*

*Bizonyítás.* Alkalmazzuk a Hölder egyenlőtlenséget mindegyik $E\xi_k^2 I(|\xi_k|>\varepsilon s_n)$, $1\le k\le n$, tagra $p=\frac{2+\alpha}{2}$ és $q=\frac{2+\alpha}{\alpha}$ választással, majd a becslésként kapott összegre alkalmazzuk megint a Hölder egyenlőtlenséget ugyanezzel a $(p,q)$ párral. Íly módon a következő becslést kapjuk.

$$\begin{aligned}\sum_{k=1}^n E\xi_k^2 I(|\xi_k|>\varepsilon s_n) &\le \sum_{k=1}^n \left(E|\xi_k|^{(2+\alpha)}\right)^{2/(2+\alpha)} P(|\xi_k|>\varepsilon s_n)^{\alpha/(2+\alpha)}\\ &\le \left(\sum_{k=1}^n E|\xi_k|^{(2+\alpha)}\right)^{2/(2+\alpha)} \left(\sum_{k=1}^n P(|\xi_k|>\varepsilon s_n)\right)^{\alpha/(2+\alpha)}.\end{aligned}$$

Másrészt, a Csebisev egyenlőtlenség alapján $\sum\limits_{k=1}^n P(|\xi_k|>\varepsilon s_n)\le \sum\limits_{k=1}^n \frac{E\xi_k^2}{\varepsilon^2 s_n^2}=\frac{1}{\varepsilon^2}$. Ezért $\left(\sum\limits_{k=1}^n P(|\xi_k|>\varepsilon s_n)\right)^{\alpha/(2+\alpha)}\le \varepsilon^{-2\alpha/(2+\alpha)}$, és ha teljesül a tétel a) feltétele, akkor

$$\limsup_{n\to\infty} \frac{1}{s_n^2}\sum_{k=1}^n E\xi_k^2 I(|\xi_k|>\varepsilon s_n)\le \varepsilon^{-2\alpha/(2+\alpha)} \lim_{n\to\infty}\left(\frac{\sum\limits_{k=1}^n E|\xi_k|^{(2+\alpha)}}{s_n^{2+\alpha}}\right)^{2/(2+\alpha)}=0$$

minden $\varepsilon>0$ számra, azaz teljesül a Lindeberg feltétel.

Az a) feltélt után megfogalmazott rész feltételeinek teljesülése esetén $s_n^2\ge \text{const.}\, n$, és $\sum\limits_{k=1}^n E|\xi_k|^{2+\alpha}=o\left(n^{(\alpha+2)/2)}\right)$, ha $n\to\infty$. Ezért ekkor teljesül az a.) részben megfogalmazott feltétel. Ha teljesül a Lindeberg feltétel, akkor teljesül a centrális határeloszlástétel is.

Értsük meg a fenti tétel szemléletes tartalmát. Tudjuk, hogy $s_n^2 = \sum\limits_{k=1}^n E\xi_k^2 = \sum\limits_{k=1}^n E\xi_k^2 I(|x_k| \le \varepsilon) + \sum\limits_{k=1}^n E\xi_k^2 I(|x_k| > \varepsilon)$ minden $\varepsilon > 0$ számra. A Lindeberg feltétel azt fejezi ki, hogy a fenti azonosság jobboldalán szereplő összeg második tagja sokkal kisebb, mint $s_n^2$. Ez szemléletesen azt jelenti, hogy az $E\xi_k^2 = \int \xi_k^2(\omega)\, dP(\omega)$ integrálokban azon $\omega$-k hozadéka, ahol $|\xi_k(\omega)|$ túl nagy, (pontosabban az $\{\omega\colon\ |\xi_k(\omega)| \ge \varepsilon s_n\}$ halmaz hozadéka) elhanyagolhatóan kicsi. Ennek kívánjuk valamilyen viszonylag könnyen ellenőrizhető, és sok érdekes esetben teljesülő feltételét megadni.

A gondolat a következő. Az $E|\xi_k|^{2+\alpha} = \int |\xi_k|^{2+\alpha}(\omega)\, dP(\omega)$ integrálokban az $\alpha > 0$ esetben azon $\omega$-k hozadéka, ahol $\xi_k(\omega)$ rendkívül nagy, nagyobb, mint az $E\xi_k^2$ kifejezést megadó integrálokban. Szép esetekben az $E|\xi_k|^{2+\alpha}$ integrál nem túl nagy. Például, ha mindegyik $\xi_k$ valószínűségi változó abszolut értéke kisebb, mint egy (a $k$ indextől nem függő) korlát, akkor $E|\xi_k|^{2+\alpha} \le \text{const.}\, E\xi_k^2$ minden $k$ indexre, és $\sum\limits_{k=1}^n E|\xi_k|^{2+\alpha} \le \text{const.}\, s_n^2$. A fenti tétel a) pontban megfogalmazott feltételében az a $\sum\limits_{k=1}^n E|\xi_k|^{2+\alpha}$ összegre egy ennél gyengébb feltételt fogalmaztunk meg. (Jegyezzük meg, hogy $s_n^2 \to \infty$ a feltételeink szerint.) Ez megengedi, hogy legyenek olyan $\omega$-k, amelyekre $\xi_k(\omega)$ viszonylag nagy, de ezek hatása a feltételek teljesülése esetén korlátozott. Sikerült olyan feltételt találni, amely biztosítja a Lindeberg feltétel teljesülését.

Annak érdekében, hogy a centrális határeloszlástétel tartalmát, illetve a Lindeberg feltétel szerepét ebben az eredményben jobban megértsük lássunk olyan példákat is, amelyekben a centrális határeloszlástétel nem érvényes, mert nem teljesül a Lindeberg feltétel. Két különböző példát fogok mutatni, amelyekben a centrális határeloszlástételben előírt függetlenségi és egyenletes kicsiségi feltételek teljesülnek, mégsem érvényes a centrális határeloszlástétel, mert nem teljesül a Lindeberg feltétel.

**Példa olyan modellre, amelyben nem teljesül a centrális határeloszlástétel.** *Legyenek $\xi_n$, $n = 1, 2, \dots$, független valószínűségi változók a következő eloszlással: $P(\xi_n = n) = P(\xi_n = -n) = \frac{1}{4n^2}$, $P(\xi_n = 1) = P(\xi_n = -1) = \frac14$, és $P(\xi_n = 0) = \frac12 - \frac{1}{2n^2}$, $n = 1, 2, \dots$. Ekkor $E\xi_n = 0$, $E\xi_n^2 = 1$. Az $S_n = \frac{1}{\sqrt{n}} \sum\limits_{k=1}^n \xi_n$ normalizált részletösszegek sorozata egyrészt teljesíti az $ES_n = 0$, és $\operatorname{Var} S_n = 1$ relációkat, másrészt, mint látni fogjuk, eloszlásban konvergál egy nulla várható értékű és $\frac12$ szórásnégyzetű normális eloszlású valószínűségi változóhoz, ha $n \to \infty$.*

Ebben a példában független valószínűségi változók olyan normalizált részletösszegeit tekintettük, amelyre nem teljesül a centrális határeloszlástétel. Ugyanis a centrális határeloszlástétel szerint az $S_n$ normalizált részletösszegeknek a nulla várható értékű és *egy szórásnégyzetű* normális eloszláshoz kellene konvergálni. Az egyenletes kicsiség feltétele teljesül ebben a példában, mert $s_n^2 = n$, és $\sup\limits_{1\le k\le n} \frac{E\xi_k^2}{s_n^2} = \frac{1}{n}$ nullához konvergál, ha $n \to \infty$. Viszont nem teljesül a Lindeberg feltétel, mert például $\varepsilon = \frac{1}{2}$ választással $\frac{1}{s_n^2} \sum\limits_{k=1}^{n} E(\xi_k^2 I(|\xi_k| > \varepsilon s_n) = \frac{1}{n} \sum\limits_{k=1}^{n} k^2 P(|\xi_k| > \frac{n}{2}) = \frac{1}{n} \sum\limits_{k=\frac{n}{2}+1}^{n} k^2 \frac{1}{2k^2} = \frac{1}{4}$.

*A példa indoklása.* Az $ES_n = 0$ és $ES_n^2 = 1$ reláció nyilvánvaló. A határeloszlásról szóló állítás bizonyítása érdekében vezessük be az $X_n = \xi_n I(|\xi_n| \le 1)$, $Y_n = \xi_n I(|\xi_n| > 1)$, $U_n = \sum\limits_{k=n}^{n} X_k$ és $V_n = \sum\limits_{k=1}^{n} Y_k$ mennyiségeket. Ekkor $S_n = \frac{1}{\sqrt{n}}(U_n + V_n)$. A független, egyforma eloszlású valószínűségi változók részletösszegeire vonatkozó centrális határeloszlástétel alapján a $\frac{1}{\sqrt{n}} U_n$, $n = 1, 2, \dots$, sorozat eloszlásban konvergál a nulla várható értékű és $\frac{1}{2}$ szórásnégyzetű normális eloszláshoz $n \to \infty$ esetén, mert az $X_k$, $EX_k = 0$, $EX_k^2 = \frac{1}{2}$, valószínűségi változók $k = 1, 2, \dots$, függetlenek és egyforma eloszlásúak. A $\frac{1}{\sqrt{n}} V_n$ kifejezések sztochasztikusan konvergálnak nullához, ha $n \to \infty$. Ugyanis $\sum\limits_{k=1}^{\infty} P(Y_k \ne 0) < \infty$, ezért a Borell–Cantelli lemma alapján egy valószínűséggel csak véges sok $Y_k(\omega)$ nem egyenlő nullával, és $\sum\limits_{k=1}^{\infty} |Y_k(\omega)| \le K(\omega)$ 1 valószínűséggel.

A példa állítása következik a fenti állításokból és az úgynevezett Slutzky lemmából. Ez utóbbi lemma azt mondja ki, hogy amennyiben valószínűségi változók valamely $P_n$ sorozata eloszlásban konvergál egy $F$ eloszláshoz, és valószínűségi változók egy másik $Q_n$ sorozata sztochasztikusan konvergál nullához, akkor a $P_n + Q_n$ sorozat is konvergál eloszlásban az $F$ eloszláshoz. E lemma bizonyítását, ami egyébként nem nehéz, itt nem tárgyalom. Egy lehetséges bizonyítás megtalálható az Első Téma jegyzet kiegészítésében. Mivel $S_n = \frac{1}{\sqrt{n}}(U_n + V_n)$, a példa állítása következik a fenti relációkból és a Slutzky lemmából $P_n = \frac{1}{\sqrt{n}} U_n$ és $Q_n = \frac{1}{\sqrt{n}} V_n$ választással.

**Egy másik példa olyan modellre, amelyben nem teljesül a centrális határeloszlástétel.** *Tekintsünk egy $\bar\xi_{k,j}$, $k = 1, 2, \dots$, $1 \le j \le n_k$, szériasorozatot, amelyre $n_k \to \infty$, ha $k \to \infty$, a $\bar\xi_{k,j}$, $1 \le j \le n_k$, valószínűségi változók függetlenek, $P(\bar\xi_{k,j} = 1) = 1 - P(\bar\xi_{k,j} = 0) = \lambda_{k,j}$, $1 \le j \le n_k$, és a $\lambda_{k,j}$ konstansok teljesítik a $\lim\limits_{k\to\infty} \sup\limits_{1\le j\le n_k} \lambda_{k,j} = 0$ és a $\lim\limits_{k\to\infty} \sum\limits_{j=1}^{n_k} \lambda_{k,j} = 1$ feltételeket. Legyen $\xi_{k,j} = \bar\xi_{k,j} - E\bar\xi_{k,j}$. A $\xi_{k,j}$, $1 \le j \le n_k$, szériasorozat a Lindeberg feltétel kivételével teljesíti a szériasorozatokra vonatkozó eloszlástétel feltételeit. Másrészt az $S_k = \sum\limits_{j=1}^{n_k} \xi_{k,j}$ sorösszegek eloszlásban konvergálnak egy $\eta - E\eta$ valószínűségi változóhoz, ahol $\eta$ Poisson eloszlású $\lambda = 1$ paraméterrel.*

*A példa indoklása.* A $\bar\xi_{k,j}$, $1 \le j \le n_k$, szériasorozat teljesíti a szériasorozatokra vonatkozó Poisson határeloszlástétel feltételeit $\lambda = 1$ paraméterrel. Ezért az $\bar S_k = \sum\limits_{j=1}^{n_k} \bar\xi_{k,j}$ összegek egy $\lambda = 1$ paraméterű Poisson eloszlású $\eta$ valószínűségi változóhoz az $S_k = \bar{S}_k - E\bar{S}_k$ valószínűségi változók pedig az $\eta - E\eta$ valószínűségi változóhoz konvergálnak eloszlásban.

Könnyen látható, hogy a $\xi_{k,j}$, $1 \le j \le n_k$ szériasorozat a Lindeberg feltételt kivéve a szériasorozatokra vonatkozó centrális határeloszlástétel összes feltételét teljesíti. A Lindeberg feltételt viszont nem teljesíti. Ugyanis létezik olyan $k = k_0$ index, hogy $E\bar{\xi}_{k,j} = \lambda_{k,j} < \frac{1}{2}$ minden $k \ge k_0$ és $1 \le j \le n_k$ indexre. Ezért $\xi^2_{k,j} I(|\xi_{k,j}| > \varepsilon) = \xi^2_{k,j} I(\bar{\xi}_{k,j} = 1)$, ha $k \ge k_0$ és $\varepsilon = \frac{1}{2}$. Innen $E\xi^2_{k,j} I(|\xi_{k,j}| > \varepsilon) = E\xi^2_{k,j} I(\bar{\xi}_{k,j} = 1) = \lambda_{k,j}(1-\lambda_{k,j})^2$, és $\sum\limits_{j=1}^{n_k} E\xi^2_{k,j} I(|\xi_{k,j}| > \varepsilon) = \sum\limits_{j=1}^{n_k} \lambda_{k,j}(1-\lambda_{k,j})^2$, ha $k \ge k_j$, és $\varepsilon = \frac{1}{2}$. Viszont nem nehéz belátni, hogy a $\lambda_{k,j}$ számokra tett feltételek mellett

$$\lim_{k\to\infty} \sum_{j=1}^{n_k} \lambda_{k,j}(1-\lambda_{k,j})^2 = 1.$$

Rátérek a centrális határeloszlástétel bizonyítására. *Az eloszlások konvergenciájáról szóló Alaptétel egyszerűsített változata* alapján elég azt bebizonyítani, hogy ha teljesülnek a megfelelő centrális határeloszlástétel feltételei, akkor a sorösszegek (szériasorozat esetén) vagy a normalizált részletösszegek (független valószínűségi változók sorozata esetén) karakterisztikus függvényei konvergálnak a standard normális eloszlás karakterisztikus függvényéhez. Ezért érdemes először a standard normális eloszlás karakterisztikus függvényét kiszámolni. Ez történik a következő lemmában. A lemma bizonyításában kihasználom a komplex függvénytan néhány fontos eredményét.

**Lemma a standard normális eloszlásfüggvény karakterisztikus függvényéről.** *A $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, $-\infty < x < \infty$, sűrűségfüggvénnyel rendelkező standard normális eloszlásfüggvény karakterisztikus függvénye a $g(t) = e^{-t^2/2}$ függvény, azaz*

$$\int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx = e^{-t^2/2}.$$

*Bizonyítás.*

$$\begin{aligned} g(t) &= \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(x-it)^2/2} e^{-t^2/2}\, dx \\ &= e^{-t^2/2} \int_{-\infty-it}^{\infty-it} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx. \end{aligned}$$

A fenti számolásban a szokásos technikát alkalmaztuk, az exponensben szereplő kvadratikus alakot teljes négyzetté alakítotuk át. Azt állítom, hogy

$$\int_{-\infty-it}^{\infty-it} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} e^{-t^2/2}\, du = 1.$$

Ez az integrál abban különbözik a standard normális sűrűségfüggvény integráljától, hogy a normális sűrűségfüggvény integrálját nem a valós tengelyen, hanem egy vele párhuzamos egyenesen tekintjük. Ez az integrál ugyanannyi, mintha a valós tengelyen integráltuk volna az $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ függvényt. Ez egyszerűen következik a komplex függvénytan talán legfontosabb eredményéből, amely szerint egy analitikus függvény körintegrálja egy zárt görbén nulla. Azt kell kihasználni, hogy a $h(z) = e^{-z^2/2}$ függvény analitikus az egész számsíkon, és ezenkívül a $h(z)$ függvény olyan, hogy amennyiben a $z$ argumentum imaginárius részének az abszolut értéke kisebb mint valamely fix $K$ szám, reális részének az abszolut értéke pedig nagyon nagy, akkor a $h(z)$ függvény nagyon kicsi. Mivel ez a bizonyítás a komplex függvénytani ismeretek felhasználása segítségével egyszerűen végrehajtható, viszont ez a komplex függvénytani rész, amelyet nem tanult mindenki elengedhetetlen ebben a bizonyításban, ezért a részletek kidolgozását elhagyom.

**Következmény.** *Két független normális eloszlású valószínűségi változó összege normális eloszlású valószínűségi változó.*

*Bizonyítás.* A következmény állítását be lehet bizonyítani két normális sűrűségfüggvény konvoluciójának a kiszámításával, és a gyakorlaton ezt meg fogjuk tenni. De egyszerűbben célhoz érünk két független normális eloszlású valószínűségi változó összegének a karakterisztikus függvényének a kiszámolásával és annak felhasználásával, hogy egy eloszlást meghatároz a karakterisztikus függvénye.

Legyen $\xi$ egy $m_1$ várható értékű és $\sigma_1^2$ szórásnégyzetű normális eloszlású valószínűségi változó. Akkor $\xi$ felírható $\xi = \sigma_1\xi_1 + m_1$ alakban, ahol $\xi_1$ standard normális eloszlású valószínűségi változó. Ezért $\xi$ karakterisztikus függvénye

$$\varphi_1(t) = Ee^{it(\sigma_1\xi_1+m_1)} = e^{itm_1}Ee^{i(t\sigma_1)\xi_1} = e^{-\sigma_1^2t^2/2+itm_1}.$$

Legyen $\eta$ a $\xi$-től független $m_2$ várható értékű és $\sigma_2^2$ szórásnégyzetű normális eloszlású valószínűségi változó. Akkor $\eta$ karakterisztikus függvénye $\varphi_2(t) = e^{-\sigma_2^2t^2/2+itm_2}$, $\xi+\eta$ karakterisztikus függvénye pedig $\varphi_1(t)\varphi_2(t) = e^{-(\sigma_1^2+\sigma_2^2)t^2/2+it(m_1+m_2)}$. Innen következik, hogy $\xi+\eta$ $m_1+m_2$ várható értékű és $\sigma_1^2+\sigma_2^2$ szórásnégyzetű normális eloszlású valószínűségi változó.

A *szériasorozatok sorösszegeire érvényes centrális határeloszlástételt* fogom bebizonyítani. A bizonyítás előtt ismertetek egy heurisztikus indoklást, majd megmutatom, hogy e heurisztikus indoklás részleteit kidolgozva hogyan kaphatunk pontos bizonyítást.

Tekintsünk egy $\xi_{k,j}$, $k = 1, 2, \dots$, $1 \le j \le n_k$, szériasorozatot, amelynek tagjaira $E\xi_{k,j} = 0$, $E\xi_{k,j}^2 < \infty$, $k = 1, 2, \dots$, $1 \le j \le n_k$, és $\lim\limits_{k\to\infty} \sum\limits_{j=1}^{n_k} E\xi_{k,j}^2 = 1$. Jelölje $\varphi_{k,j}(t) = Ee^{it\xi_{k,j}}$, $-\infty < t < \infty$, a $\xi_{k,j}$ valószínűségi változó karakterisztikus függvényét, és vezessük be az $S_k = \sum\limits_{j=1}^{n_k} \xi_{k,j}$ jelölést.

Ekkor azt kell megmutatnunk, hogy $\lim\limits_{k\to\infty} Ee^{itS_k} = e^{-t^2/2}$ minden $-\infty < t < \infty$ számra. Viszont tudjuk a független valószínűségi változók szorzatának várható értékéről szóló eredmény alapján, hogy

$$Ee^{itS_k} = E\exp\left\{\sum_{j=1}^{n_k} it\xi_{k,j}\right\} = \prod_{j=1}^{n_k} Ee^{it\xi_{k,j}} = \prod_{j=1}^{n_k} \varphi_{k,j}(t) \tag{1}$$

A fenti azonosságban logaritmust véve a bizonyítandó állítást így fogalmazhatjuk át:

$$\lim_{k\to\infty} \sum_{j=1}^{n_k} \log\varphi_{k,j}(t) = -t^2/2, \quad -\infty < t < \infty. \tag{2}$$

Viszont a $\varphi_{k,j}(t)$ karakterisztikus függvényeket Taylor sorba fejtve a $t=0$ pont körül, és felhasználva a *valószínűségi változók momentumainak kifejezéséről a karakterisztikus függvényük segítségével megadott kifejezéséről* szóló eredményt az előző előadásból a $\varphi_{k,j}(t)$ karakterisztikus függvényekre a következő közelítést tudjuk adni. (Itt felhasználjuk, hogy mint azt később látni fogjuk, a tekintett valószínűségi változók kicsisége miatt az alábbi Taylor sorfejtéseket a nulla egy kis környezetében végezzük.)

$$\varphi_{k,j}(t) \sim \varphi_{k,j}(0) + t\varphi'_{k,j}(0) + \frac{t^2}{2}\varphi''_{k,j}(0) = 1 + itE\xi_{k,j} - \frac{t^2}{2}E\xi^2_{k,j} = 1 - \frac{t^2}{2}E\xi^2_{k,j}.$$

Továbbá, mivel $\log(1-u) \sim u$ kis $u$ számokra, azt várjuk, hogy a $\log\varphi_{k,j}(t) \sim -\frac{t^2}{2}E\xi^2_{k,j}$ jó közelítést ad. E közelítő relációkat összegezve, és felhasználva a tétel feltételeit azt kapjuk, hogy $\sum\limits_{j=1}^{n_k} \log\varphi_{k,j}(t) \sim -\frac{t^2}{2}\sum\limits_{j=1}^{n_k} E\xi^2_{k,j} \sim -\frac{t^2}{2}$, és ezt kellett megmutatni. Belátjuk, hogy a fenti közelítő azonosságok pontosabb kifejtése segítségével be tudjuk bizonyítani a kívánt centrális határeloszlástételt. E program végrehajtásának érdekében azonban fel kell tenni, hogy teljesül a Lindeberg feltétel.

A részletek kidolgozása előtt belátok egy a bizonyításban hasznos lemmát, amely arról szól, hogy az $e^{iu}$ trigonometrikus függvényt milyen jól közelítik Taylor sorának szeletei.

**Lemma trigonometrikus függvények Taylor sor közelítéséről.** *Tetszőleges nem negatív egész $k$ és valós $u$ számra*

$$\left| e^{iu} - \left(1 + \frac{iu}{1!} + \cdots + \frac{(iu)^k}{k!}\right)\right| \le \frac{|u|^{k+1}}{(k+1)!}. \tag{3}$$

*Bizonyítás:* Vezessük be az $F(u) = e^{iu} - \left(1 + \frac{iu}{1!} + \cdots + \frac{(iu)^k}{k!}\right)$ függvényt, és tekintsük ennek $F^{(j)}(u)$ deriváltjait. $F^{(j)}(0) = 0$, ha $0 \le j \le k$, és $|F^{(k+1)}(u)| = |e^{iku}| = 1$ minden $u$ valós számra. Teljes indukcióval kapjuk, hogy $|F^{(j)}(u)| \le \int_0^u |F^{(j+1)}(s)|\,ds \le$

$\int_0^u \frac{|s|^{k-j}\,ds}{(k-j)!} = \frac{|u|^{k+1-j}}{(k+1-j)!}$ minden $j=k+1,k,\dots,0$ számra. Tehát $|F(u)|\le\frac{|u|^{k+1}}{(k+1)!}$, és ez a Lemma állítása.

*A centrális határeloszlástétel bizonyítása szériasorozatok sorösszegeire.* Lássuk először be, hogy a Lindeberg feltétel teljesüléséből következik az egyenletes kicsiség tulajdonsága.

Rögzítsünk egy $\varepsilon>0$ számot. Ekkor

$$E\xi_{k,j}^2 = E\xi_{k,j}^2 I(\{(|\xi_{k,j}|<\varepsilon\}) + E\xi_{k,j}^2 I(\{|\xi_{k,j}|\ge\varepsilon\}) \le \varepsilon^2 + \sum_{j=1}^{n_k} E\xi_{k,j}^2 I(\{|\xi_{k,j}|\ge\varepsilon\}),$$

ezért a Lindeberg feltétel alapján $\limsup\limits_{k\to\infty}\sup\limits_{1\le j\le n_k} E\xi_{k,j}^2\le\varepsilon^2$. Mivel ez az állítás minden $\varepsilon>0$ számra érvényes, innen következik az egyenletes kicsiség feltétele.

A centrális határeloszlástétel bizonyítása érdekében először azt mutatom meg, hogy az $E\xi_{k,j}=0$ reláció és a Lindeberg feltétel teljesülése miatt

$$\lim_{k\to\infty}\sum_{j=1}^{n_k}(\varphi_{k,j}(t)-1) = \lim_{k\to\infty}\sum_{k=1}^{n_k} E\left(e^{it\xi_{k,j}}-1-it\xi_{k,j}\right) = -\frac{t^2}{2}. \tag{4}$$

Mivel $\lim\limits_{k\to\infty}\sum\limits_{j=1}^{n_k} E\xi_{k,j}^2=1$, ez ekvivalens azzal, hogy

$$\lim_{k\to\infty}\sum_{j=1}^{n_k} E\left(e^{it\xi_{k,j}}-1-it\xi_{k,j}+\frac{t^2}{2}\xi_{k,j}^2\right)=0.$$

Alkalmazva a (3) formulát $u=t\xi_{k,j}$ választással $k=2$-re, ha $|t\xi_{k,j}|\le\varepsilon$ és $k=1$-re, ha $|t\xi_{k,j}|\ge\varepsilon$ azt kapjuk, hogy

$$\begin{aligned}\left|\sum_{j=1}^{n_k} E\left(e^{it\xi_{k,j}}-1-it\xi_{k,j}+\frac{t^2}{2}\xi_{k,j}^2\right) I(\{|\xi_{k,j}|\le\varepsilon\})\right| &\le \sum_{j=1}^{n_k} E\frac{|t\xi_{k,j}|^3}{6} I(\{|\xi_{k,j}|\le\varepsilon\})\\ &\le \varepsilon|t|^3\sum_{j=1}^{n_k} E\frac{\xi_{k,j}^2}{6}\le \text{const.}\,\varepsilon,\end{aligned}$$

és a Lindeberg feltétel alapján

$$\begin{aligned}&\lim_{k\to\infty}\left|\sum_{j=1}^{n_k} E\left(e^{it\xi_{k,j}}-1-it\xi_{k,j}+\frac{t^2}{2}\xi_{k,j}^2\right))I(\{|\xi_{k,j}|>\varepsilon\})\right|\\ &\qquad\le \lim_{k\to\infty}\sum_{j=1}^{n_k} Et^2\xi_{k,j}^2 I(\{|\xi_{k,j}|>\varepsilon\})=0.\end{aligned}$$

Mivel ezek a relációk minden $\varepsilon = 0$-ra érvényesek, innen következik a (4) formula.

Meg akarom mutatni, hogy igaz a (4) formula azon módosítása, amelyben az $1 - \varphi_{k,j}(t)$ függvényt a $\log\varphi_{k,j}(t)$ függvénnyel helyettesítem, azaz igaz a (2) formulában felírt azonosság. Ennek érdekében belátom a következő technikai jellegű lemmát.

**Lemma karakterisztikus függvény logaritmusának a becsléséről.** *Tekintsük egy $\xi$ valószínűségi változó $\varphi(t)$ karkakterisztikus függvényét valamilyen rögzített $t$ számra. Ha $E\xi = 0$, $E\xi^2 \le \varepsilon$ egy elég kis $\varepsilon = \varepsilon(t) > 0$ számmal akkor $|1 - \varphi(t)| \le \frac{t^2}{2}E\xi^2$, és $|\log\varphi(t) + (1 - \varphi(t))| \le \frac{t^4}{4}\left(E\xi^2\right)^2$.*

*A lemma bizonyítása.* Alkalmazva a (3) formulát $k = 1$-re kapjuk, hogy $|e^{it\xi} - 1 - it\xi| \le \frac{t^2\xi^2}{2}$. Ezért véve a baloldalon az abszolútérték jelek közötti kifejezés várható értékét kapjuk, hogy $|\varphi(t) - 1| \le \frac{t^2}{2}E\xi^2$. Ha $E\xi^2 < \varepsilon$ elég kis $\varepsilon = \varepsilon(t) > 0$ számmal, akkor $|1 - \varphi(t)| \le \frac{1}{4}$. Vegyük észre, hogy véve a $\log(1 - z)$ függvény Taylor sorát a $|z| \le \frac{1}{4}$ halmazon, azt kapjuk, hogy $|\log(1 - z) - z| \le |z^2|$. Innen $|\log\varphi(t) + (1 - \varphi(t))| = |\log\left(1 - (1 - \varphi(t))\right) - (1 - \varphi(t))| \le |1 - \varphi(t)|^2 \le \frac{t^4}{4}\left(E\xi^2\right)^2$, ha $E\xi^2 \le \varepsilon(t)$.

*Megjegyzés.* A fenti bizonyításban is felhasználtam a komplex függvénytan néhány fontos eredményét. Ugyanis $\varphi(t)$ komplex értékű függvény, így a számolásban a logaritmus függvény analitikus kiterjesztésével kellett dolgoznunk. Azt az eredményt használtam ki, hogy az $e^{1+z}$ függvénynek kis $z$, pontosabban $|z| < \frac{1}{4}$, számokra létezik olyan egyértelmű $\log(1 + z)$ inverze, amely a $z = 0$ pontban nullával egyenlő, és ezt a függvényt ki lehet fejezni a $\log(1 + z) = \sum\limits_{k=1}^{\infty} \frac{(-1)^{k+1}z^k}{k}$ hatványsor segítségével.

A bizonyítandó tétel feltételeinek teljesülése esetén, a $\xi_{k,j}$, $k = 1, 2, \dots$, $1 \le j \le n_k$, szériasorozat teljesíti az egyenletes kicsiség tulajdonságát, ezért alkalmazhatjuk a fenti lemmát mindegyik $\xi_{k,j}$, $1 \le j \le n_k$ valószínűségi változóra elég nagy (a $t$ paramétertől függő) $k$ indexre. Innen kapjuk, hogy

$$\limsup_{k\to\infty} \left| \sum_{j=1}^{n_k} \log\varphi_{k,j}(t) + \sum_{j=1}^{n_k}(1 - \varphi_{k,j}(t)) \right| \le \limsup_{k\to\infty} \frac{t^4}{4} \sum_{j=1}^{n_k} \left(E\xi_{k,j}^2\right)^2 \le \varepsilon\frac{t^4}{4}.$$

Az utolsó egyenlőtlenség azért igaz, mert $\sup\limits_{1\le j\le n_k} E\xi_{k,j}^2 \le \varepsilon$, ha $k \ge k_0(\varepsilon)$ valamely $k_0$ küszöbindex-szel, és $\lim\limits_{k\to\infty} \sum\limits_{j=1}^{n_k} E\xi_{k,j}^2 = 1$. Mivel a fenti egyenlőtlenség minden $\varepsilon > 0$ számra érvényes, ezért

$$\lim_{k\to\infty} \left| \sum_{j=1}^{n_k} \log\varphi_{k,j}(t) - \sum_{j=1}^{n_k}(\varphi_{k,j}(t) - 1) \right| = 0.$$

Ebből a relációból, és a (4) formulából következik a (2) formula. Ezután véve mind a két oldal exponenciális hatványát a (2) formulában, és felhasználva az (1) formula azonosságát azt kapjuk, hogy $\lim\limits_{k\to\infty} Ee^{itS_k} = e^{-t^2/2}$. Ezt az azonosságot akartuk bebizonyítani.

Röviden tárgyalni fogom az úgynevezett lokális centrális határeloszlástételt független, egyforma és rácsos eloszlású valószínűségi változók összegére. Megfogalmazom a fő eredményt, és teszek néhány megjegyzést, amelyek segítenek megérteni az eredmény tartalmát, illetve azt, hogy miért természetes várni egy ilyen eredményt. A tétel részletes bizonyítását a kiegészítésben írom le.

Az eredmény megfogalmazása előtt bevezetem a rácsos eloszlás fogalmát, illetve definiálom azt, hogy mikor nevezhetjük az egész számok halmazát egy eloszlást tartalmazó legritkább rácsnak.

**Rácsos eloszlás fogalma.** *Egy $\xi$ valószínűségi változó rácsos eloszlású az egész számok rácsán, ha értékei az egész számok rácsára vannak koncentrálva, azaz egy valószínűséggel csak egész értékeket vesz fel, másképp fogalmazva $\sum\limits_{k=-\infty}^{\infty} P(\xi = k) = 1$. Azt mondjuk, hogy az egész számok rácsa a $\xi$ valószínűségi változó eloszlását tartalmazó legritkább rács, ha tetszőleges $A > 1$ és $B$ egész számokra $\sum\limits_{k=-\infty}^{\infty} P(\xi = Ak + B) < 1$.*

*Általánosabban, egy $\xi$ valószínűségi változót rácsos eloszlásúnak nevezünk, ha azok értékei egy valószínűséggel egy $\{b + kh\colon\ k = 0, \pm1, \pm2, \dots\}$ alakú halmazra vannak koncentrálva valamilyen $h > 0$ és $b$ valós számokkal, azaz $\sum\limits_{k=-\infty}^{\infty} P(\xi = kh + b) = 1$. Azt mondjuk, hogy a $\xi$ valószínűségi változó értékei egy $h$, $h > 0$, szélességű rácsra (mint legritkább rácsra) vannak koncentrálva, ha létezik olyan $b$ valós szám, amelyre $\sum\limits_{k=-\infty}^{\infty} P(\xi = kh + b) = 1$, és $\sum\limits_{k=-\infty}^{\infty} P(\xi = Akh + B) < 1$ tetszőleges $A > 1$ egész és $B$ valós számokra.*

Megfogalmazom a következő, az egész számok halmazára, mint legritkább rácsra koncentrált eloszlású független és egyforma eloszlású valószínűségi változók összegére vonatkozó lokális centrális határeloszlástételnek nevezett eredményt.

**Lokális centrális határeloszlástétel független, egyforma, rácsos eloszlású valószínűségi változók összegére.** *Legyen $\xi_1, \xi_2, \dots,$ független, egyforma eloszlású valószínűségi változók sorozata, mely valószínűségi változók rácsos eloszlásúak az egész számok rácsán, és az egész számok rácsa a legritkább a $\xi_j$ valószínűségi változók (közös) eloszlását tartalmazó rács. Legyen $E\xi_1 = m$, $E\xi_1^2 = m_2 < \infty$, (tehát feltesszük, hogy a $\xi_1$ valószínűségi változó második momentuma véges), és legyen $\sigma^2 = m_2 - m_1^2$ a $\xi_1$ valószínűségi változó szórásnégyzete. Tekintsük az $S_n = \xi_1 + \dots + \xi_n$, $n = 1, 2, \dots,$ részletösszegeket. Ekkor*

$$P(S_n = k) = \frac{1}{\sqrt{2\pi n}\sigma} \exp\left\{-\frac{(k - nm)^2}{2n\sigma^2}\right\} + o\left(\frac{1}{\sqrt{n}}\right), \quad k = 0, \pm1, \pm2, \dots,$$

*ahol $o(\cdot)$ egyenletes a $k$ változóban.*

*Megjegyzés.* Legyen $\xi_1, \xi_2, \dots$ független, egyforma eloszlású valószínűségi változók sorozata, amelyek rácsos eloszlásúak, és eloszlásuk egy $kh+b$ $k=0,\pm1,\pm2,\dots$, $h$ szélességű rácsra, mint legritkább rácsra, legyen koncentrálva. Vezessük be az $S_n = \sum\limits_{j=1}^{n} \xi_j$, $n = 1,2,\dots$, részletösszegeket. A fenti tétel segítségével jó aszimptotikus becslést tudunk adni a $P(S_n = kh + nb)$ alakú valószínűségekre is. Ennek érdekében vezessük be a $\bar\xi_j = \frac{\xi_j - b}{h}$, $j = 1,2,\dots$, az egész számok rácsára (mint legritkább rácsra) koncentrált valószínűségi változókat, és ezek $\bar S_n = \sum\limits_{j=1}^{n} \bar\xi_j$ részletösszegeit. Ekkor a $P(S_n = kh+b) = P(\bar S_n = k)$ azonosság segítségével az előző tétel jó aszimptotikát ad a minket érdeklő valószínűségre.

Értsük meg a fenti lokális centrális határeloszlástétel tartalmát, és azt is, hogy miért nevezik ezt az eredményt *lokális* centrális határeloszlástételnek.

A centrális határeloszlástétel szerint $P\left(\frac{S_n - nm}{\sqrt{n}\sigma} < x\right) \sim \Phi(x)$, ahol $\Phi(\cdot)$ a standard normális eloszlásfüggvény. Ezenkívül azt is tudjuk, hogy a most vizsgált esetben az $\frac{S_n - nm}{\sqrt{n}\sigma}$ valószínűségi változó a $\frac{k-nm}{\sqrt{n}\sigma}$, $k = 0,\pm1,\pm2,\dots$, $\frac{1}{\sqrt{n}\sigma}$ sűrűségű rácson veszi fel az értékeit. Azt várjuk, hogy annak a valószínűsége, hogy az $\frac{S_n - nm}{\sqrt{n}\sigma}$ valószínűségi változó a $[\frac{k-nm-\frac12}{\sqrt{n}\sigma}, \frac{k-nm+\frac12}{\sqrt{n}\sigma}]$ intervallumban veszi fel az értékét, azaz $S_n = k$, körülbelül annyi, mint annak a valószínűsége, hogy egy standard normális eloszlású valószínűségi változó is ebben a kis intervallumban veszi fel az értékét. Ennek valószínűsége pedig $\Phi\left(\frac{k-nm-\frac12}{\sqrt{n}\sigma}\right) - \Phi\left(\frac{k-nm+\frac12}{\sqrt{n}\sigma}\right) \sim \frac{1}{\sqrt{2\pi n}\sigma} \exp\left\{-\frac{(k-nm)^2}{2n\sigma^2}\right\}$. A lokális centrális határeloszlástétel ezt az állítást mondja ki pontosabb formában. A lokális jelző a tételben arra utal, hogy a $P(S_n = k)$ valószínűség viselkedése az $S_n$ eloszlás lokális viselkedését írja le.

A lokális centrális határeloszlástétel bizonyításának az az alapgondolata, hogy a $P(S_n = k)$ valószínűségeket pontosan ki tudjuk számolni a $\xi_1$ valószínűségi változó $\varphi(t) = Ee^{it\xi_1} = \sum\limits_{k=-\infty}^{\infty} P(\xi_1 = k)e^{itk}$ karakterisztikus függvénye segítségével. Valóban, $Ee^{itS_n} = \varphi^n(t)$, és a Fourier sorok együtthatóit kifejező formula segítségével azt kapjuk, hogy

$$P(S_n = k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ikt} \varphi^n(t)\, dt. \tag{5}$$

A feladat ezután az, hogy adjunk jó aszimptotikus formulát az (5) formula jobboldalán szereplő integrálra. Ennek érdekében érdemes először az integrálban szereplő $\varphi^n(t)$ kifejezésre jó becslést adni. Tudjuk, hogy $\varphi(0) = 1$, $\varphi'(0) = imt$, $\varphi''(0) = -m_2$. Ezért azt várjuk, hogy kis $t$ értékekre a karakterisztikus függvény origó körüli Taylor sorfejtése alapján elég jó közelítést kapunk a $\varphi(t)$ függvényre, illetve annak $n$-ik hatványára. A karakterisztikus függvény Taylor sorfejtése azt sugallja, hogy $\varphi^n(t) \sim \left(1 + imt - \frac{m_2 t^2}{2}\right)^n \sim e^{n(imt - m_2 t^2/2)}$ elég jó közelítés kis $t$ argumentumokra. Valójában érdemes a fenti érvelés kissé finomított változatát alkalmazni. Célszerű az alkalmas közelítés kiszámításában nem a $\varphi(t)$, hanem a $\log\varphi(t)$ Taylor sor közelítésével dolgozni és utána a $\varphi^n(t) = e^{n\log\varphi(t)}$ azonosságot használni. De ez az eljárás is csak kis $t$ paraméterekre ad jó közelítést a $\varphi(t)^n$ mennyiségre, és felmerül a kérdés, hogy hogyan tudjuk jól becsülni a karakterisztikus függvényt nem kis $t$ argumentumokra.

Ezen a ponton jelenik meg az a feltétel, hogy a $\xi_1$ valószínűségi változó eloszlásását tartalmazó legritkább rács az egész számok rácsa. Arra, hogy a $\xi_1$ valószínűségi változó rácsos eloszlásású azért volt szükség, mert ez biztosítja az (5) formula érvényességét. A következő lemmában belátjuk, hogy a $\xi_1$ eloszlásának előbb említett tulajdonságából következik, hogy a $[-\pi,\pi]$ intervallumban, azaz az (5) formulában szereplő integrál integrálási tartományában a $t=0$ pont az egyetlen olyan pont, ahol $|\varphi(t)|=1$. Ennek a ténynek, illetve a karakterisztikus függvény folytonosságának a segítségével könnyen belátható, hogy az $\varepsilon\le|t|\le\pi$ tartomány hozadéka exponenciálisan kicsi az (5) formulában szereplő integrálban minden $\varepsilon>0$ számra.

**Lemma egy rácsos eloszlású valószínűségi változó karakterisztikus függvényének a viselkedéséről.** *Legyen egy $\xi$ valószínűségi változó eloszlása az egész számok rácsára, (mint legritkább rácsra) koncentrálva. Tekintsük a $\xi$ valószínűségi változó karakterisztikus függvényét definiáló $P(t)=\sum\limits_{k=-\infty}^{\infty} e^{ikt}P(\xi=k)$ Fourier sort. A $P(t)$ Fourier sor periódusa $2\pi$, $P(0)=1$, $|P(t)|\le 1$ minden valós $t$ számra, és $|P(t)|<1$, ha $|t|\le\pi$, és $t\neq 0$.*

*A lemma bizonyítása.* Nyílvánvaló módon $|P(t)|\le 1$ minden $t$ valós számra, és $P(0)=1$. Mutassuk meg, hogy a $\xi$ valószínűségi változó eloszlásának egy $h>0$ szám akkor és csak akkor periódusa, azaz akkor és csak akkor koncentrálódik $\xi$ eloszlása valamely $kh+b$, $k=0,\pm1,\pm2,\dots$ alakú halmazra, ha $|Ee^{i2\pi\xi/h}|=1$. Valóban, ha $|Ee^{i2\pi\xi/h}|=1$, akkor $Ee^{i2\pi\xi/h}=e^{i2\pi a}$, azaz $Ee^{i2\pi(\xi/h-a)}=1$ valamilyen, valós $a$ számra. Mivel $\operatorname{Re} e^{i2\pi x}\le 1$ minden $x$ valós számra, és $\operatorname{Re} e^{i2\pi x}=1$ csak akkor teljesül, ha $x$ egész szám, ez az azonosság csak akkor állhat fenn, ha a $\frac{\xi}{h}-a$ valószínűségi változó eloszlása a $k$, $k=0,\pm1,\pm2,\dots$, pontokba van koncentrálva, azaz a $\xi$ valószínűségi változó a $kh+b$, $k=0,\pm1,\pm2,\dots$, rácsra van koncentrálva, ahol $b=ha$. Megfordítva, ha a $\xi$ valószínűségi változó eloszlása egy $kh+b$, $k=0,\pm1,\dots$, rácsra van koncentrálva, akkor $|Ee^{i2\pi\xi}|=|e^{i2\pi b}|=1$.

Ha a $\xi$ valószínűségi változó eloszlása nincsen semmilyen $h>1$ szélességű rácsra koncentrálva, akkor az előbbiek alapján $|P(\frac{2\pi}{h})|<1$ minden $h>1$ számra, tehát a lemma feltételeinek teljesülése esetén $|P(t)|<1$ minden $0<t<2\pi$, speciálisan minden $0<t\le\pi$ számra. Mivel $P(-t)=\overline{P(t)}$, ahol a felülvonás komplex konjugáltat jelöl, ezért $|P(t)|<1$, ha $|t|\le\pi$, és $t\neq 0$. A lemmát bebizonyítottuk.

Mivel minden karakterisztikus függvény folytonos, a fenti lemmából következik, hogy a *független, egyforma, rácsos eloszlású valószínűségi változók összegére megfogalmazott lokális centrális határeloszlástétel* feltételeinek teljesülése esetén minden $\varepsilon>0$ számhoz létezik olyan $\delta=\delta(\varepsilon)>0$ szám, amelyre a $\xi_1$ valószínűségi változó $\varphi(t)$ karakterisztikus függvénye teljesíti a $|\varphi(t)|\le(1-\delta)$ egyenlőtlenséget minden $\varepsilon\le|t|\le\pi$

számra. Ezért minden $\varepsilon > 0$ számhoz létezik olyan $0 < \delta < 1$ szám, amelyre

$$\left| \frac{1}{2\pi} \int_{\{t\colon\, \varepsilon \le |t| \le \pi\}} e^{-ikt} \varphi^n(t)\, dt \right| \le (1-\delta)^n. \tag{6}$$

Ez azt jelenti, hogy az (5) formula jobboldalán szereplő integrál becslésében elég az integrálnak egy $[-\varepsilon, \varepsilon]$ intervallumba eső részére jó becslést adni. Ez, mint jeleztem, lehetséges a karakterisztikus függvény origó körüli Taylor sorfejtése segítségével. A részletes bizonyítást a kiegészítésben adom meg.

Megjegyzem, hogy a Stirling formula bizonyításában hasonló érvelést alkalmaztunk a Poisson eloszlásra. Az egyetlen különbség az volt, hogy ott a karakterisztikus függvényre egyszerű, explicit formulát tudtunk adni, ami egyszerűsítette a számolásokat.

A rácsos eloszlású valószínűségi változókra bizonyított lokális centrális határeloszlástételt a Fourier sorok Fourier együtthatóit kifejező (5) formula segítségével tudjuk bizonyítani. Ismertetek egy az (5) formulához hasonló eredményt, amely lehetővé teszi egy függvény értékeinek kiszámítását a függvény Fourier transzformáltjának segítségével.

**Inverziós formula egy függvény Fourier transzformáljára.** *Legyen $f(x)$ integrálható függvény a számegyenesen, azaz tegyük fel, hogy $\int_{-\infty}^{\infty} |f(x)|\, dx < \infty$. Tekintsük az $f(\cdot)$ függvény Fourier transzformáltját, azaz az $\tilde f(t) = \int_{-\infty}^{\infty} e^{itx} f(x)\, dx$, $-\infty < x < \infty$, függvényt. Ha az $\tilde f(t)$ függvény szintén integrálható, azaz $\int_{-\infty}^{\infty} |\tilde f(t)|\, dt < \infty$, akkor $f(x)$ a Lebesgue mérték szerint majdnem minden $x \in R^1$ pontra megegyezik az alábbi folytonos és korlátos függvénnyel:*

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \tilde f(t)\, dt. \tag{7}$$

Ezen inverziós formula segítségével sűrűségfüggvénnyel rendelkező független valószínűségi változók normalizált összegének a sűrűségfüggvényére is lehet lokális centrális határeloszlástételt bizonyítani. Ez azt állítja, hogy független, sűrűségfüggvénnyel rendelkező valószínűségi változók normalizált összegének a (létező) sűrűségfüggvénye alkalmas feltételek teljesülése esetén minden $-\infty < x < \infty$ pontban konvergál az $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ standard normális sűrűségfüggvényhez, és ez a konvergencia egyenletes. Az eredmény pontos megfogalmazását és bizonyítását, amely a rácsos eloszlású valószínűségi változók összegéről szóló centrális határeloszlástétel bizonyításához hasonló, és a (7) formulában felírt integrál becslésén alapul, elhagyom.

**Néhány a jegyzetben megfogalmazott eredmény bizonyítása.**

Először a következő eredmény bizonyítását ismertetem.

*Szériasorozatok sorösszegeiről szóló centrális határeloszlástétel megfordításának bizonyítása.* Jelölje $\varphi_{k,j}(t)$ a $\xi_{k,j}$ valószínűségi változó karakterisztikus függvényét. Először azt mutatom meg, hogy a tétel feltételeinek teljesülése esetén

$$\lim_{k\to\infty}\sum_{j=1}^{n_k}\operatorname{Re}(\varphi_{k,j}(t)-1)=-\frac{t^2}{2}. \tag{8}$$

Valóban a centrális határeloszlástételből, illetve abból a kissé gyengébb állításból, hogy a $\sum\limits_{j=1}^{n_k}\xi_{k,j}$ sorösszegek eloszlásban konvergálnak egy 1 szórásnégyzetű (ismeretlen $m$ várható értékű) normális valószínűségi változóhoz következik, hogy

$$\lim_{k\to\infty}\prod_{j=1}^{n_k}\varphi_{k,j}(t)=e^{-t^2/2+imt}\quad\text{minden valós } t \text{ számra,}$$

illetve először abszolut értéket majd logaritmust véve

$$\lim_{k\to\infty}\sum_{j=1}^{n_k}\operatorname{Re}(\log\varphi_{k,j}(t))=-\frac{t^2}{2}\quad\text{minden valós } t \text{ számra.} \tag{9}$$

Továbbá az egyenletes kicsiség feltétele miatt minden $\varepsilon>0$ számhoz létezik olyan $k_0=k_0(\varepsilon)$ küszöbindex, amelyre $E\xi_{k,j}^2\le\varepsilon$ minden $1\le j\le n_k$ indexre, ha $k\le k_0$. Ezért *a karakterisztikus függvény logaritmusának a becsléséről* szóló lemma alapján minden $t>0$ és elég kis $\varepsilon=\varepsilon(t)>0$ számhoz létezik olyan $k_0=k_0(\varepsilon)$ küszöbindex, amelyre $|\operatorname{Re}(\log\varphi_{k,j}(t)+\operatorname{Re}(1-\varphi_{k,j}(t)))|\le|\log\varphi_{k,j}(t)+(1-\varphi_{k,j}(t))|\le\frac{t^4}{4}\left(E\xi_{k,j}^2\right)^2\le\frac{\varepsilon t^4}{4}E\xi_{k,j}^2$ minden $k\ge k_0$ és $1\le j\le n_k$ indexre. Ezen egyenlőtlenségeket összegezve, és felhasználva a $\lim\limits_{k\to\infty}\sum\limits_{j=1}^{n_k}E\xi_{k,j}^2=1$ relációt, azt kapjuk, hogy

$$\limsup_{k\to\infty}\left|\sum_{j=1}^{n_k}\operatorname{Re}(\log\varphi_{k,j}(t))+\sum_{j=1}^{n_k}\operatorname{Re}(1-\varphi_{k,j}(t))\right|\le\frac{\varepsilon t^4}{4}.$$

Mivel ez a reláció minden $\varepsilon>0$ számra igaz, ezért a (9) relációval együtt implikálja a (8) formulát.

Mivel nagy $k$ indexre a szériasorozat sorainak az összege közel egy szórásnégyzetű, ezért a (8) formulából következik, hogy

$$\lim_{k\to\infty}\sum_{j=1}^{n_k}E\left(\cos(t\xi_{k,j})-1+\frac{t^2\xi_{k,j}^2}{2}\right)=0\quad\text{minden } t\in R^1 \text{ számra.} \tag{10}$$

Vegyük észre, hogy $\cos u - 1 + \frac{u^2}{2} \ge 0$ minden $u \in R^1$ számra. Valóban, az $F(u) = \cos u - 1 + \frac{u^2}{2}$ függvényre $F''(u) = 1 - \cos u \ge 0$ minden $u$ valós számra, és $F(0) = F'(0) = 0$. Innen az $F'(u) = \int_0^u F''(s)\,ds$ függvényre $F'(u) \ge 0$, ha $u \ge 0$, és $F'(u) \le 0$, ha $u < 0$. Hasonlóan, további integrálással $F(u) \ge 0$ minden $u$ számra. Ezenkívül $\cos u - 1 + \frac{u^2}{2} \ge -2 + \frac{u^2}{2} \ge \frac{u^2}{4}$, ha $|u| > 3$. Ezért a (10) formulában szereplő várható értékekre felírhatjuk, hogy

$$E\left(\cos(t\xi_{k,j}) - 1 + \frac{t^2\xi_{k,j}^2}{2}\right) = E\left(\cos(t\xi_{k,j}) - 1 + \frac{t^2\xi_{k,j}^2}{2}\right) I(|t\xi_{k,j}| \le 3)$$
$$+ E\left(\cos(t\xi_{k,j}) - 1 + \frac{t^2\xi_{k,j}^2}{2}\right) I(|t\xi_{k,j}| > 3) \ge \frac{t^2}{4} E\xi_{k,j}^2 I(|t\xi_{k,j}| > 3).$$

Ezt felhasználva a (10) formulából következik, hogy

$$\lim_{k\to\infty} \sum_{j=1}^{n_k} \frac{t^2}{4} E\xi_{k,j}^2 I\left(\left\{|\xi_{k,j}| \ge \frac{3}{t}\right\}\right) = 0,$$

azaz

$$\lim_{k\to\infty} \sum_{j=1}^{n_k} E\xi_{k,j}^2 I\left(\left\{|\xi_{k,j}| \ge \frac{3}{t}\right\}\right) = 0$$

minden $t$ számra. Innen $t = \frac{3}{\varepsilon}$ választással megkapjuk a tétel állítását.

Ismertetem még a következő bizonyítást is.

*A lokális centrális határeloszlástétel bizonyítása független, egyforma, rácsos eloszlású valószínűségi változók összegére.* Az (5) formula alapján

$$P(S_n = k) = \int_{-\pi}^{\pi} \frac{1}{2\pi} e^{-ikt} P^n(t)\,dt = \int_{|t|<\frac{1}{\varepsilon\sqrt{n}}} + \int_{\frac{1}{\varepsilon\sqrt{n}}<|t|<\varepsilon} + \int_{\varepsilon<|t|<\pi} = I_1 + I_2 + I_3,$$

ahol $P(t) = \sum\limits_{k=-\infty}^{\infty} e^{ikt} P(\xi_1 = k)$, és $\varepsilon > 0$ tetszőleges kis pozitív szám. A tétel bizonyítása érdekében jó becslést adunk az $I_1$, $I_2$ és $I_3$ integrálokra.

Az $I_3$ integrált már megbecsültük a (6) formulában. Az $I_1$ és $I_2$ integrálok kiszámításához jó becslést kell adnunk a $P^n(t)$ függvényre, ha $|t| < \varepsilon$. Kényelmesebb a $\log P(t)$ függvénnyel dolgozni. (Ez kis $\varepsilon > 0$ számra lehetséges, mert ebben az esetben a $P(t)$ függvény értéke a $[-\varepsilon, \varepsilon]$ intervallumban szeparálva van nullától.) Egyszerű számolással kapjuk, hogy

$$\frac{d\log P(t)}{dt} = \frac{P'(t)}{P(t)}, \qquad \left.\frac{d\log P(t)}{dt}\right|_{t=0} = im,$$
$$\frac{d^2\log P(t)}{dt^2} = \frac{P''(t)P(t) - P'(t)^2}{P^2(t)}, \qquad \left.\frac{d^2\log P(t)}{dt^2}\right|_{t=0} = -m_2 + m^2 = -\sigma^2,$$

ezért az origó körüli Taylor sorfejtéssel kapjuk, hogy

$$\log P(t) = imt - \frac{\sigma^2}{2}t^2 + o(t^2), \quad \text{ha } |t| < \varepsilon.$$

Innen, mivel $|P^n(t)| = e^{n\mathrm{Re}\,\log P(t)} = e^{-n(\sigma^2t^2/2+o(t^2))} \le e^{-n\sigma^2t^2/3}$, ha $|t| < \varepsilon$ és $n \ge n(\varepsilon)$, ezért

$$\begin{aligned} |I_2| &\le \frac{1}{2\pi}\int_{\frac{1}{\varepsilon\sqrt{n}}<|t|<\varepsilon} |P^n(t)|\,dt \le \frac{1}{2\pi}\int_{\frac{1}{\varepsilon\sqrt{n}}<|t|<\varepsilon} e^{-n\sigma^2t^2/3}\,dt \\ &\le \frac{1}{\pi\sqrt{n}}\int_{-\frac{1}{\varepsilon}}^{\infty} e^{-\sigma^2t^2/3}\,dt \le e^{-\sigma^2/4\varepsilon^2}. \end{aligned}$$

Továbbá

$$\begin{aligned} I_1 &= \int_{-\frac{1}{\varepsilon\sqrt{n}}}^{\frac{1}{\varepsilon\sqrt{n}}} \frac{1}{2\pi}e^{-ikt+inmt-n\sigma^2t^2/2+o(nt^2)}\,dt \\ &= \int_{-\frac{1}{\varepsilon}}^{\frac{1}{\varepsilon}} \frac{1}{2\pi\sqrt{n}}e^{i(mn-k)t/\sqrt{n}-\sigma^2t^2/2+o(t^2)}\,dt \\ &= \int_{-\infty}^{\infty} \frac{1}{2\pi\sqrt{n}}e^{i(nm-k)t/\sqrt{n}-\sigma^2t^2/2}\,dt - \int_{|t|>\frac{1}{\varepsilon}} \frac{1}{2\pi\sqrt{n}}e^{i(nm-k)t/\sqrt{n}-\sigma^2t^2/2}\,dt \\ &\qquad + o\left(\frac{1}{\sqrt{n}}\right). \end{aligned}$$

Másrészt

$$\left|\int_{|t|>\frac{1}{\varepsilon}} \frac{1}{2\pi\sqrt{n}}e^{i(nm-k)t/\sqrt{n}-\sigma^2t^2/2}\,dt\right| \le e^{-\sigma^2/4\varepsilon^2}$$

és az alábbi integrál exponensében szereplő kvadratikus alakot kiegészítve teljes négyzetté azt kapjuk, hogy

$$\begin{aligned} &\int_{-\infty}^{\infty} \frac{1}{2\pi\sqrt{n}}e^{i(nm-k)t/\sqrt{n}-\sigma^2t^2/2}\,dt \\ &\quad = \frac{e^{-(nm-k)^2/2n\sigma^2}}{2\pi\sqrt{n}}\int_{-\infty}^{\infty} \exp\left\{-\frac{\sigma^2}{2}\left(t - i\frac{(nm-k)}{\sqrt{n}\sigma^2}\right)^2\right\}dt = \frac{e^{-(nm-k)^2/2n\sigma^2}}{\sqrt{2\pi n}\sigma} \end{aligned}$$

a komplex függvénytan integrálási szabályai szerint. Egyébként az utolsó azonosság a normális eloszlás karakterisztikus függvényét kifejező képlet eredménye alapján is látható. Ezekből a becslésekből következik, hogy

$$\left|I_1 - \frac{1}{\sqrt{2\pi n}\sigma}\exp\left\{-\frac{(k-nm)^2}{2n\sigma^2}\right\}\right| \le \text{const.}\, e^{-\sigma^2/4\varepsilon^2},$$

ha $n > n(\varepsilon)$. Mivel az $I_1$, $I_2$ és $I_3$ kifejezésekre adott becslések tetszőlegesen kicsi $\varepsilon > 0$-ra érvényesek, ha $n = n(\varepsilon)$ elég nagy, innen következik a tétel állítása.

**Néhány további kiegészítő megjegyzés.**

Tárgyalok néhány olyan problémát és eredményt, amelyek természetes módon megjelennek a centrális határeloszlástétel vizsgálata során. Ezen eredmények többségét nem fogalmazom meg pontosan, és megelégszem azok érvényességének egy heurisztikus indoklásával.

Láttuk, hogy a Fourier sorok Fourier együtthatóit kifejező formula, illetve az inverz Fourier transzformációt kifejező formula segítségével úgynevezett lokális centrális eloszlástételt tudunk bizonyítani független, rácsos eloszlású vagy sűrűségfüggvénnyel rendelkező egyforma eloszlású valószínűségi változók összegére. Az ilyen jellegű tételek azt mondják ki, hogy megfelelő feltételek teljesülése esetén független valószínűségi változók alkalmasan normalizált összegeinek sűrűségfüggvénye konvergál a standard normális eloszlás sűrűségfüggvényéhez. A bizonyítás azon alapul, hogy a keresett sűrűségfüggvényt ki tudjuk fejezni egy alkalmas integrál segítségével, amelynek a magfüggvényében szerepel a karakterisztikus függvény $n$-ik hatványának alkalmasan átskálázott alakja. Némi számolással be lehet látni, hogy úgynevezett szinguláris integrálokat kell vizsgálni, amelyeknek a lényeges hozadéka a nulla kis környezetébe van koncentrálva, és az integrál maradék része elhanyagolhatóan kicsi. A karakterisztikus függvény, pontosabban a karakterisztikus függvény logaritmusának kis értékeit jól megbecsülhetjük e függvény origó körüli második tagig vett Taylor sor fejtésének a segítségével, és ez a közelítés lehetővé teszi a sűrűségfüggvényt kifejező integrál olyan jó aszimptotikus becslését, amelyből következik a centrális határeloszlástétel.

Természetes gondolat, hogy a karakterisztikus függvény logaritmusának több tagot tartalmazó Taylor sorát véve pontosabb becslést lehet adni e függvényre, és ezáltal független valószínűségi változók normalizált összegének a sűrűségfüggvényére is. Ez a program végrehajtható, és ilyen módon alkalmas feltételek teljesülése esetén meg lehet adni ezen normalizált összeg sűrűségfüggvényének az úgynevezett Edgeworth sorfejtését. Ez egy olyan véges összeg, amelynek az első, azaz fő tagja a standard normális sűrűségfüggvény. Ezt követi egy korrekciós tag, amelyben $n^{-1/2}$-szer a standard normális sűrűségfüggvény van megszorozva egy alkalmas polinommal. Ezt a közelítést esetleg még tovább finomíthatjuk egy további korrekciós taggal, amelyben $n^{-1}$-szer a standard normális sűrűségfüggvény van megszorozva egy alkalmas polinommal. Ezt az eljárást tovább is folytathatjuk. A $k$-ik korrekciós tag $n^{-k/2}$-szer a standard normális sűrűségfüggvény megszorozva egy alkalmas polinommal. Az első $k$ korrekciós tagot figyelembe vevő közelítés $n^{-(k+1)/2}$ pontossággal közelíti a sűrűségfüggvényt, a benne szereplő polinomok explicit módon kiszámolhatóak, és azon véletlen összeg tagjainak az első $k+2$ momentumától függnek, amelynek a normalizált sűrűségfüggvényét vizsgáljuk.

Felmerül a kérdés, hogy létezik-e hasonló Edgeworth sorfejtés független valószínűségi változók összegének az eloszlásfüggvényére is. (Tehát nemcsak a sűrűség, hanem az eloszlásfüggvényre is szeretnénk alkalmas sorfejtést találni.) Ilyen eredményt is be lehet bizonyítani, bár ez a probléma nehezebb. A fő nehézség abban rejlik, hogy az eloszlásfüggvényt nem tudjuk a sűrűségfüggvényhez hasonló viszonylag egyszerű formulával kifejezni a karakterisztikus függvény segítségével. Ezen a problémán alkalmas simítási eljárások segítségével lehet segíteni. Ilyen simítási eljárások segítségével az eloszlásfüggvények vizsgálatát vissza lehet vezetni a sűrűségfüggvények vizsgálatára.

A fent említett kérdések és eredmények elsősorban elvi szempontból érdekesek, gyakorlati problémákban ritkán jelennek meg. Fontosabb a következő kérdés vizsgálata. Tekintsük független, egyforma eloszlású $\xi_1, \xi_2, \dots,$ valószínűségi változók $S_n = \sum\limits_{k=1}^{n} \xi_k$, $n = 1, 2, \dots,$ részletösszegeit. A centrális határeloszlástétel jó aszimptotikus formulát ad a $P\left(\frac{S_n - nE\xi_1}{\sqrt{n \operatorname{Var} \xi_1}} > x\right)$ valószínűségekre rögzített $x$ számra, ha $n \to \infty$. Tudunk-e hasonló aszimptotikát adni a $P\left(\frac{S_n - nE\xi_1}{\sqrt{n \operatorname{Var} \xi_1}} > x_n\right)$ valószínűségekre akkor, ha $x_n \to \infty$, azaz az $n$ paraméterrel együtt az $x_n$ szám is végtelenhez tart? Különösen fontos az az eset, amikor $x_n = x\sqrt{n}$ valamely rögzített $x > 0$ számmal. Ez azt jelenti, hogy a $P\left(\frac{S_n - nE\xi_1}{n} > x\sqrt{\operatorname{Var} \xi_1}\right)$ valószínűség értékére vagyunk kiváncsiak. A nagy számok törvénye azt mondja ki, hogy ez a valószínűség nullához tart, ha $n \to \infty$. De szeretnénk tudni e konvergencia pontos nagyságrendjét is. Erre a kérdésre kielégítő válasz ismeretes, és ezt a választ a nagy eltérés tételnek nevezik. Megfogalmazom ezt a tételt.

**Nagy eltérések tétele független, egyforma eloszlású valószínűségi változók összegére.** *Legyen $\xi_1, \xi_2, \dots$ független, egyforma eloszlású valószínűségi változók sorozata, amelyekre $E\xi_1 = 0$, és $R(t) = Ee^{t\xi_1} < \infty$ valamely $t > 0$ számmal. Jelölje $S_n = \sum\limits_{k=1}^{n} \xi_k$, $n = 1, 2, \dots,$ e valószínűségi változók részletösszegeit. Ezek teljesítik a következő aszimptotikus relációt.*

$$\lim_{n\to\infty} -\frac{1}{n} \log P\left(\frac{S_n}{n} > x\right) = \rho(x) \quad \textit{minden } x > 0 \textit{ számra,}$$

*ahol $\rho(x) = \sup\limits_{t\colon t \ge 0} (tx - \log R(t))$, és $R(t) = Ee^{t\xi_1}$, ha $Ee^{t\xi_1} < \infty$, és $R(t) = \infty$ egyébként. $\rho(x) > 0$ minden $x > 0$ számra, és $\rho(x) = \frac{x^2}{2 \operatorname{Var} \xi_1} + O(x^3)$ kis $x > 0$ számokra.*

A fenti tétel azt mondja ki, hogy rögzített $x > 0$ számra a $P\left(\frac{S_n}{n} > x\right)$ valószínűség exponenciálisan kicsi az $n$ változóban, és megadja az exponens nagyságrendjét. Ezt a nagyságrendet a $\rho(x)$ függvény adja meg, amit az $R(t)$ függvény Legendre transzformáltjának neveznek az irodalomban. Kis $x$ számokra ez a nagyságrend körülbelül annyi, mint amennyit a centrális határeloszlástétel sugall. Ezt fejezi ki a $\rho(x) = \frac{x^2}{2 \operatorname{Var} \xi_1} + O(x^3)$ aszimptotikus azonosság. De ez csak közelítőleg igaz. Be lehet látni, hogy a $\rho(x)$ függvény értékei a nulla egy tetszőleges kis környezetében egyértelműen meghatározzák a $\xi_1$ valószínűségi változó eloszlását. Megjegyzem azt is, hogy a nagy eltérés tétel eredménye lényegesen különbözik az Edgeworth sorfejtésben kapott eredménytől. Ott a $P(S_n > \sqrt{n}x)$ valószínűségre olyan aszimptotikát kaptunk, amelynek a hibatagja $O(n^{-(k+1)/2}$ valamilyen rögzített $k$ egész számmal. Ilyen pontosságú becslések semmit sem mondanak olyan esetben, amikor exponenciálisan kicsi valószínűségeket becslünk.

A nagy eltérés tétel feltételei között szerepelt az $Ee^{t\xi_1} < \infty$ egyenlőtlenség valamely $t > 0$ paraméterrel. Érdemes ennek a feltételnek a szerepét is megérteni. Emlékeztetek arra, hogy a klasszikus centrális határeloszlástétel feltételei között szerepelt a Lindeberg feltétel, ami egy olyan jellegű megkötés volt, hogy a tekintett valószínűségi változók csak kis valószínűséggel lehetnek nagyon nagyok. Az $Ee^{t\xi_1} < \infty$ feltétel egy hasonló jellegű, csak erősebb megkötést ír elő. Valóban, a Markov egyenlőtlenség alapján e feltételből következik, hogy $P(\xi_1 > x) = P(e^{t\xi_1} > e^{tx}) \le \frac{Ee^{t\xi_1}}{e^{tx}} = \text{const.}\, e^{-tx}$, azaz a $\xi_1$ valószínűségi változó csak exponenciálisan kis valószínűséggel vesz fel nagy $x$ értékeket. A következő példa célja annak megvilágítása, hogy miért van szükség a nagy eltérés tételben egy ilyen feltételre.

Tekintsük független, egyforma eloszlású $\xi_1, \xi_2, \dots,$ valószínűségi változók olyan sorozatát, amelyben a $\xi_1$ valószínűségi változónak $f(x) = \text{const.}\, e^{-|x|^\alpha}$ alakú sűrűségfüggvénye van valamely $0 < \alpha < 1$ paraméterrel. (A const. együtthatót úgy választjuk, hogy $f(x)$ sűrűségfüggvény legyen.) Ekkor $E\xi_1 = 0$, (mert az $f(x)$ sűrűségfüggvény páros), $E|\xi_1|^k < \infty$ minden $k > 0$ számra, azaz $\xi_1$ minden momentuma véges, de $Ee^{t\xi_1} = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx \ge \text{const.} \int_0^\infty e^{tx - x^\alpha}\,dx = \infty$ minden $t > 0$ számra, azaz a nagy eltérés tétel feltételei nem teljesülnek. Legyen $S_n = \sum\limits_{k=1}^n \xi_k$. Azt állítom, hogy $P\left(\frac{S_n}{n} > x\right) \ge e^{-\text{const.}\, n^\alpha}$ minden $x > 0$ számra valamely az $x$ számtól függő konstanssal, azaz az ebben a példában tekintett véletlen összeg nem teljesíti a nagy eltérések tétel becslését. A $P\left(\frac{S_n}{n} > x\right)$ valószínűség túl nagy, nem exponenciálisan kicsi.

Valóban, párossági meggondolások alapján $P\left(\sum\limits_{k=2}^n \xi_k \ge 0\right) = \frac{1}{2}$, és

$$\begin{aligned} P\left(\frac{S_n}{n} > x\right) &\ge P\left(\xi_1 > nx,\ \frac{1}{n}\sum_{k=2}^n \xi_k \ge 0\right) = P(\xi_1 > nx) P\left(\frac{1}{n}\sum_{k=2}^n \xi_k \ge 0\right) \\ &= \frac{1}{2} P(\xi_1 > nx) = \frac{1}{2}\int_{nx}^\infty f(x)\,dx \ge e^{-\text{const.}\, n^\alpha} \end{aligned}$$

alkalmas konstanssal, amint állítottuk.

A lokális centrális határeloszlástétel, illetve annak finomításaiban a sűrűségfüggvényre a Fourier analízis segítségével felírt integrálok aszimptotikájára adtunk jó becslést. Az eloszlásfüggvények aszimptotikáját megadó centrális határeloszlástétel bizonyításának a hátterében is ezek a formulák rejtőznek. Megpróbálhatnánk ezen formulák alaposabb vizsgálatával a nagy eltérések tételét is bebizonyítani. Egy ilyen módszer azonban nem működik jól. A problémát az okozza, hogy olyan integrálokat kell jól becsülni, amelyekben a (komplex szám értékű) integrandus nagyon erősen oszcillál. Az ilyen integrálok becslésére a komplex függvénytanban kidolgozták az úgynevezett nyeregpont módszert, és ennek alkalmazásával lehet a nagy eltérés tételt is bebizonyítani. Az $R(t) = Ee^{t\xi_1} < \infty$ feltételre azért van szükség a bizonyításban, mert ez teszi lehetővé a komplex függvénytani módszerek alkalmazását. A részletek kidolgozásától eltekintek.