A április 17-i előadáshoz kapcsolódó gyakorlat feladatai

Először néhány példát mutatunk a centrális határeloszlástétel alkalmazására.

1. Egy szabályos dobókockát feldobunk 1200 alkalommal egymástól függetlenül, és összeadjuk a páros értékű dobások eredményét. Adjunk jó közelítő becslést a centrális határeloszlástétel és egy normális eloszlástáblázat segítségével arra, hogy ez az összeg 2280 és 2500 közé esik.

   Megoldás: Vezessük be a következő ξ_j, $1 \le j \le 1200$, valószínűségi változókat: $\xi_j = 2$, ha a j-ik dobás eredménye 2, $\xi_j = 4$, ha a j-ik dobás eredménye 4, $\xi_j = 6$, ha a j-ik dobás eredménye 6, $\xi_j = 0$, ha a j-ik dobás eredménye 1, 3 vagy 5. Ekkor a $P\left(2280 \le \sum\limits_{j=1}^{1200} \xi_j \le 2500\right)$ valószínűséget kell jól megbecsülnünk. Vegyük észre, hogy $E\xi_j = \frac{1}{6}(2+4+6) = 2$, $\mathrm{Var}\,\xi_j = \frac{1}{6}(4+16+36) - 4 = \frac{16}{3}$. Innen a centrális határeloszlástétel alapján

$$P\left(2280 \le \sum_{j=1}^{1200} \xi_j \le 2500\right) = P\left(\frac{-120}{\sqrt{1200\frac{16}{3}}} \le \frac{\sum\limits_{j=1}^{1200} \xi_j - \sum\limits_{j=1}^{1200} E\xi_j}{\sqrt{\sum\limits_{j=1}^{1200} \mathrm{Var}\,\xi_j}} \le \frac{100}{\sqrt{1200\frac{16}{3}}}\right)$$
$$\sim \Phi(1.25) - \Phi(-1.5) = 0.8944 + 0.9322 - 1 = 0.8266.$$

2. 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 1830 küzé esik? Adjunk erre a valószínűségre jó közelítő becslést.

   Megoldás: Az elvégzett dobások száma egy η negatív binomiális eloszlású valószínűségi változó $n = 1200$ és $p = \frac{2}{3}$ paraméterekkel, azaz $P(\eta = k + n) = \binom{n+k-1}{n-1}(1-p)^k p^n$, $p = \frac{2}{3}$, és $n = 1200$ paraméterrel. Egy ilyen valószínűségi változónak ki lehet számolni a pontos eloszlását, azaz azt, hogy milyen értéket milyen valószínűséggel vesz fel. Elvileg, ez lehetőséget ad a kívánt valószínűség kiszámítására egy bonyolult összeg kiszámításának a segítségével. 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 $n = 1$, $p = \frac{2}{3}$ paraméterekkel, és minket a $P(1680 < \xi_1 + \cdots + \xi_{1200} < 1830)$ valószínűség érdekel. Megmutattuk a 6. előadásban, illetve a hozzátartozó feladatokban), hogy $E\xi_j = \frac{1}{p} = \frac{3}{2}$, $\mathrm{Var}\,\xi_j = \frac{1-p}{p^2} = \frac{3}{4}$. Ezért a centrális határeloszlástétel alapján $\eta = \xi_1 + \cdots + \xi_{1200}$ jelöléssel minket a $P\left(-4 < \frac{\eta - 1200E\xi_1}{\sqrt{1200\mathrm{Var}\,\xi_1}} < 1\right)$ valószínűség érdekel. A centrális határeloszlástétel alapján $P\left(-4 < \frac{\eta - 1200E\xi_1}{\sqrt{1200\mathrm{Var}\,\xi_1}} < 1\right) \sim \Phi(1) + \Phi(4) - 1 \sim \Phi(1)$.

3. Legyen birtokunkban 100 lámpa, melyek mindegyike egymástól független időtartamig működik, élettartamuk pedig exponenciális eloszlású $\lambda = \frac{1}{10}$ 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.

   Megoldás: 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.

   Kiszámoltuk, hogy jelen esetben $E\eta = mE\xi_1 = \frac{m}{\lambda} = 1000$, $\mathrm{Var}\,\eta = \frac{m}{\lambda^2} = 10000$ ($m = 100$ és $\lambda = \frac{1}{10}$ választással). Ezért a centrális határeloszlástétel szerint $\frac{\eta - E\eta}{\sqrt{\mathrm{Var}\,\eta}} = \frac{\eta - 1000}{100}$ jó közelítéssel standard normális eloszlású valószínűségi változó, és $P(\xi_1 + \cdots + \xi_{100} > 1150) = P\left(\frac{\eta - E\eta}{\sqrt{\mathrm{Var}\,\eta}} > 1.5\right) \sim 1 - \Phi(1.5)$.

4. Egy pénzdarabról ellenőrizni akarjuk, hogy igaz-e az a hipotézis, mely szerint ez az érme legalább $\frac{3}{4}$ valószínűséggel esik a fej és legfeljebb $\frac{1}{4}$ valószínűséggel az írás oldalára. Ennek érdekében feldobjuk a pénzdarabot 30 000 alkalommal, és a következő döntési szabályt hozzuk. Választunk egy k számot, és akkor fogadjuk el a hipotézist helyesnek, ha legalább k fejdobás történt. Legalább mekkorának kell válassztanunk ezt a k számot, ha azt akarjuk, hogy egy a hipotézist teljesítő pénzdarab esetén legalább 0.9 valószínűséggel döntsünk úgy, hogy a hipotézis teljesül?

   Megoldás: Vezessük be a következő valószínűségi változókat: $\xi_j = 1$, ha a j-ik dobás eredménye fej, $\xi_j = 0$, ha a j-ik dobás eredménye írás, $1 \le j \le 30\,000$, $S = S_{30000} = \sum\limits_{j=1}^{30\,000} \xi_j$. Ha a fejdobás eredményének valószínűsége pontosan $\frac{3}{4}$, $E\xi_j = \frac{3}{4}$, $E\xi_j^2 = \frac{3}{4}$, $\mathrm{Var}\,\xi_j = E\xi_j^2 - (E\xi_j)^2 = \frac{3}{16}$, $ES = 30\,000E\xi_j = 22\,500$, $\mathrm{Var}\,S = 30\,000\mathrm{Var}\,\xi_j = 5625 = 75^2$. Innen és a centrális határeloszlástételből,

$$P(S > k) = P\left(\frac{S - ES}{\sqrt{\mathrm{Var}\,S}} > \frac{k - 22\,500}{75}\right) = 1 - P\left(\frac{S - ES}{\sqrt{\mathrm{Var}\,S}} \le \frac{k - 22\,500}{75}\right)$$

$$\sim 1 - \Phi\left(\frac{k - 22\,500}{75}\right).$$

Válasszuk a k számot úgy, hogy a fenti valószínűség körülbelül 0.9 legyen. Ekkor a $\Phi\left(\frac{k-22\,500}{75}\right) = 0.1$ vagy ami ezzel ekvivalens, a $\Phi\left(\frac{22\,500-k}{75}\right) = 0.9$ egyenletet kell kielégítenünk. A normális eloszlás-táblázat alapján $\frac{22\,500-k}{75} \sim 1.28$, ami azt jelenti, hogy $k = 22\,500 - 75 \times 1.28$ és $p = \frac{3}{4}$ esetén annak valószínűsége, hogy a fejdobások száma nagyobb mint $k = 22\,500 - 75 \times 1.28 = 22\,212$ és $p = \frac{3}{4}$ esetében annak valószínűsége, hogy legalább ennyi fejdobás történik körülbelül 0.9. Ha $p \ge \frac{3}{4}$, akkor ez a valószínűség nagyobb. Ezért a $k = 22\,212$ helyes választás.

Házi feladat.

Egy szabályos dobókockát feldobunk 2700 alkalommal egymástól függetlenül, és összeszámoljuk a páros értékű dobások eredményét. Adjunk jó közelítő becslést a centrális határeloszlástétel és egy normális eloszlástáblázat segítségével arra, hogy ez az összeg 420 és 720 közé esik.

A következő két feladat célja annak az alaptételnek nevezett tétel feltételeinek jobb megértése, mely megadja a kapcsolatot az eloszlásfüggvények illetve azok karakterisztikus függvénye konvergenciája közötti kapcsolatot.

Ezenkívül, ha marad rá idő meg lehet tárgyalni egy ennek a feladatsornak a végén megfogalmazott állítás bizonyítását, mely megmagyarázza mi a jelentősége az alaptételben annak a feltételnek, hogy a karakterisztikus függvények egy a nullában folytonos függvényhez konvergálnak. Egyébként ennek az állításnak a bizonyítása az alaptétel bizonyításának egyik kulcslépése.

5. Mutassuk meg, hogy egy F eloszlásfüggvény $\varphi(t) = \int e^{itu} F(\,du)$ karakterisztikus függvénye folytonos minden pontban.
Megoldás: Jelölje $\varphi(t) = \int e^{itu} F(\,du)$ az F eloszlásfüggvény karakterisztikus függvényét. Ekkor tetszőleges t, δ illetve $K > 0$ számra teljesül a

$$|\varphi(t+\delta) - \varphi(t)| \le \int |e^{i(t+\delta)} - e^{it}| F(\,du) \le \int_{u:\ |u|\le K} \sup_{t:\ |t|\le K} |e^{i(t+\delta)} - e^{it}| F(\,du)$$
$$+ \int_{u:\ |u|>K} 2F(\,du) \le \sup_{t:\ |t|\le K} |e^{i\delta t} - 1| + F(-K) + (1 - F(K))$$

egyenlőtlenség. Tetszőleges $\varepsilon > 0$ számhoz választhatunk olyan $K = K(\varepsilon)$ számot, melyre $F(-K) + (1 - F(K)) \le \frac{\varepsilon}{2}$. Ezután elég kis $\delta_0 = \delta_0(\varepsilon, K, t) > 0$ választással elérhetjük, hogy $\sup_{t:\ |t|\le K} |e^{i\delta t} - 1| \le \frac{\varepsilon}{2}$, ha $\delta \le \delta_0$. Innen következik, hogy $|\varphi(t + \delta) - \varphi(t)| \le \varepsilon$, ha $\delta < \delta_0$, azaz a $\varphi(\cdot)$ függvény folytonos.

6. Legyen $F_n(\cdot)$, $n = 1, 2 \dots$, a $[-n, n]$ intervallumban egyenletes eloszlású valószínűségi változók sorozata, azaz legyen $F_n(\cdot)$ sűrűségfüggvénye az $f_n(u) = \frac{1}{2n}$, ha $-n \le u \le n$, $f_n(u) = 0$, ha $|u| > n$ képlet segítségével megadott függvény. Lássuk be, hogy az $F_n(\cdot)$ eloszlásfüggvények karakterisztikus függvényei minden pontban konvergálnak az $\varphi(0) = 1$, $\varphi(t) = 0$, ha $t \ne 0$ képlettel megadott $\varphi(t)$ függvényhez, amelyik az origóban nem folytonos. Továbbá az $F_n(\cdot)$ eloszlásfüggvények nem konvergálnak eloszlásban, mert minden véges $[a, b]$ intervallumra

$$\lim_{n\to\infty} [F_n(b) - F_n(a)] = 0.$$

Megoldás: Az F_n eloszlásfüggvény karakterisztikus függvénye

$$\varphi_n(t) = \frac{1}{2n} \int_{-n}^{n} e^{itu}\, du = \frac{e^{itn} - e^{-itn}}{2nit},$$

ahonnan $\varphi_n(t) \le \frac{1}{|t|n}$. Innen $\lim_{n\to\infty} \varphi_n(t) = 0$, ha $t \ne 0$. Másrészt, $\varphi_n(0) = 1$ minden n-re tehát, $\lim_{n\to\infty} \varphi_n(0) = 1$. Másrészt egy véges $[a, b]$ intervallumra az $F_n(b) - F_n(a) = \frac{b-a}{n}$, ha $n \le \max(|a|, |b|)$. ahonnan $\lim_{n\to\infty} [F_n(b) - F_n(a)] = 0$. Ez azt jelenti, hogy az F_n eloszlások „kifolynak a végtelenbe", és ezért nincs határeloszlásuk.

7. Mutassuk meg, a karakterisztikus függvények segítségével, hogy amennyiben ξ és η két független, normális eloszlású valószínűségi változó m_1 és m_2 várható értékkel, σ_1^2 és σ_2^2 szórásnégyzettel, akkor $\xi + \eta$ normális eloszlású valószínűségi változó $m_1 + m_2$ várható értékkel, és $\sigma_1^2 + \sigma_2^2$ szórásnégyzettel.

   Megoldás: Tudjuk, hogy a karakteresztikus függvények egyértelműen meghatározzák azt az eloszlásfüggvényt, melynek ők a karakterisztikus függvényük. Ezért elég belátni, hogy a $\xi + \eta$ karakterisztikus függvénye megegyezik az $m_1 + m_2$ várható értékkel és $\sigma_1^2 + \sigma_2^2$ szórásnégyzettel rendelkező normális eloszlás karakterisztikus függvényével. Viszont tudjuk egyrészt, hogy a ξ és η valószínűségi változók függetlensége miatt $Ee^{it(\xi+\eta)} = Ee^{it\xi}Ee^{it\eta}$, másrészt mivel a standard normális eloszlásfüggvény karakterisztikus függvénye $\varphi(t) = e^{-t^2/2}$, ezért $Ee^{it\xi} = e^{itm_1}e^{-\sigma_1^2 t^2/2}$, és $Ee^{it\eta} = e^{itm_2}e^{-\sigma_2^2 t^2/2}$. Innen $Ee^{it(\xi+\eta)} = e^{it(m_1+m_2)}e^{-(\sigma_1^2+\sigma_2^2)t^2/2}$, ahonnan következik a feladat állítása.

8. Legyenek ξ és η független valószínűségi változók, melyek közül ξ Poisson eloszlású λ paraméterrel, azaz $P(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}$ minden $k = 0, 1, 2, \dots$ számra, η egyenletes eloszlású a $[0, 1]$ intervallumon, azaz létezik $f(u)$ sűrűségfüggvénye, melyre $f(u) = 1$, ha $0 \le u \le 1$, és $f(u) = 0$, ha $u > 1$, vagy $u < 0$. Lássuk be, hogy a $\xi + \eta$ valószínűségi változónak is van sűrűségfüggvénye, és határozzuk meg azt.

   Megoldás: Legyen $[a, b]$ egy olyan intervallum, melyre $[a, b] \subset [n, n + 1]$ valamely nem negatív egész n számra. Ekkor

$$\begin{aligned} P(\xi+\eta\in[a,b]) &= P(\xi=n,\eta\in[a-n,b-n]) = P(\xi=n)P(\eta\in[a-n,b-n]) \\ &= \frac{\lambda^n}{n!}e^{-\lambda}(b-a) = \int_a^b f(u)\,du, \end{aligned}$$

ahol $f(u)=\dfrac{\lambda^n}{n!}e^{-\lambda}$, ha $u\in[n,n+1]$, $n=0,1,2,\dots$. Legyen továbbá $f(u)=0$, ha $u\le 0$. Ebből a relációból, illetve abból a tényből, hogy $P(\xi+\eta\le 0)=0$, következik, hogy $P(\xi+\eta<x)=\int_{-\infty}^x f(u)\,du$ a fent definiált $f(\cdot)$ függvénnyel, tehát ez a $\xi+\eta$ valószínűségi változó (létező) sűrűségfüggvénye.

9. Legyenek ξ és η független valószínűségi változók, melyek közül ξ Poisson eloszlású λ paraméterrel, azaz $P(\xi=k)=\dfrac{\lambda^k}{k!}e^{-\lambda}$ minden $k=0,1,2,\dots$ számra, η egyenletes eloszlású a $[0,2]$ intervallumon, azaz létezik $f(u)$ sűrűségfüggvénye, melyre $f(u)=\dfrac12$, ha $0\le u\le 2$, és $f(u)=0$, ha $u>2$, vagy $u<0$. Lássuk be, hogy a $\xi+\eta$ valószínűségi változónak is van sűrűségfüggvénye, és határozzuk meg azt.
Megoldás: Ennek a feladatnak a megoldása hasonló az előzőhöz. Legyen $[a,b]$ egy olyan intervallum, melyre $[a,b]\subset[n,n+1]$ valamely nem negatív egész n számra. Ekkor, ha $n\ge 1$, akkor

$$\begin{aligned} P(\xi+\eta\in[a,b]) &= P(\xi=n,\eta\in[a-n,b-n]) \\ &\qquad + P(\xi=n-1,\eta\in[a-n-1,b-n-1]) \\ &= P(\xi=n)P(\eta\in[a-n,b-n]) \\ &\qquad + P(\xi=n-1)P(\eta\in[a-n-1,b-n-1]) \\ &= \left(\frac{\lambda^n}{n!}+\frac{\lambda^{n-1}}{(n-1)!}\right)e^{-\lambda}(b-a) = \int_a^b f(u)\,du, \end{aligned}$$

ahol $f(u)=\left(\dfrac{\lambda^n}{n!}+\dfrac{\lambda^{n-1}}{(n-1)!}\right)e^{-\lambda}$, ha $u\in[n,n+1]$, $n=1,2,\dots$. Hasonlóan,

$$P(\xi+\eta\in[a,b]) = \int_a^b P(\xi=0)\,du = \int_a^b e^{-\lambda}\,du$$

ha $[a,b]\in= 0,1]$, továbbá $P(\xi+\eta\le 0)=0$. Innen következik, hogy ha az $f(\cdot)$ függvény definicióját az $f(u)=e^{-\lambda}$, ha $0\le u\le 1$, $f(u)=0$, ha $u<0$ képletekkel fejezzük be, akkor $P(\xi+\eta<x)=\int_{-\infty}^x f(u)\,du$ ezzel az $f(\cdot)$ függvénnyel. Tehát ez a $\xi+\eta$ valószínűségi változó (létező) sűrűségfüggvénye a fent definiált $f(\cdot)$ függvény.

10. Legyen $F_n(\cdot)$, $n=1,2,\dots$, eloszlásfüggvények egy sorozata, melyek $\varphi_n(\cdot)$ karakterisztikus függvényei az origó egy kis környezetében konvergálnak egy a nullában folytonos függvényhez.
    a.) Lássuk be (a Lebesgue tétel segítségével), hogy ekkor tetszőleges $\varepsilon>0$ számhoz létezik olyan $\delta=\delta(\varepsilon)$ szám, melyre $\lim\limits_{n\to\infty}\dfrac{1}{2\delta}\displaystyle\int_{-\delta}^{\delta}\Re[1-\varphi_n(t)]\,dt\le\varepsilon$, ahol $\Re z$ a z szám reális részét jelöli.

b.) Mutassuk meg, hogy amennyiben $F_n(\cdot)$ eloszlásfüggvények $\varphi_n(\cdot)$ karakterisztikus függvényei teljesítik az a) részben megfogalmazott tulajdonságot, akkor minden $\varepsilon > 0$ számhoz létezik olyan $K = K(\varepsilon)$ szám, melyre $F_n(-K) + (1 - F_n(K)) \le \varepsilon$ minden $n = 1, 2, \dots$ számra.

Megjegyzés: A b.) részben megfogalmazott tulajdonság teljesülését egyenletes feszességnek hívják az irodalomban. Ez fontos szerepet játszik számos vizsgálatban.

Megoldás: Legyen $\varphi(\cdot)$ a $\varphi_n(\cdot)$ függvények (egy véges intervallumban létező) limesze. Mivel ez a függvény az origóban folytonos $0 \le \Re\varphi(t) \le 1$ minden t indexre, ahol a $\varphi(\cdot)$ létezik, $\Re[1-\varphi(0)] = 0$, és ez a függvény az origóban folytonos, ezért $\frac{1}{2\delta}\int_{-\delta}^{\delta} \Re[1-\varphi(t)]\,dt \le \varepsilon$, ha a $\delta = \delta(\varepsilon) > 0$ számot elég kicsinek választjuk. Mivel $\lim\limits_{n\to\infty} \Re[1-\varphi_n(t)] = \Re[1-\varphi(t)]$, és $0 \le \Re[1-\varphi_n(t)] \le 2$, alkalmazhatjuk a Lebesgue tételt, és ez megadja az a) rész állításának bizonyítását.

A b) rész bizonyításához használjuk fel, hogy az a) rész teljesülése esetén minden $\varepsilon > 0$ számhoz választhatunk olyan $\delta = \delta(\varepsilon)$ számot és $n_0 = n_0(\varepsilon)$ küszöbindexet, melyre igaz, hogy $0 \le \frac{1}{2\delta}\int_{-\delta}^{\delta} \Re[1-\varphi_n(t)]\,dt \le \frac{\varepsilon}{2}$, ha $n \ge n_0$. Másrészt

$$\begin{aligned}
\frac{1}{2\delta}\int_{-\delta}^{\delta} \Re[1-\varphi_n(t)]\,dt &= \int_{-\delta}^{\delta} \frac{1}{2\delta}\int_{-\infty}^{\infty} [1-\cos tx]\; dF_n(x)\,dt \\
&= \int_{-\infty}^{\infty} \frac{1}{2\delta}\int_{-\delta}^{\delta} [1-\cos tx]\,dt\,dF_n(x) = \int_{-\infty}^{\infty} \left[\frac{t}{2\delta} - \frac{\sin tx}{2\delta x}\right]_{t=-\delta}^{t=\delta} dF_n(x) \\
&= \int_{-\infty}^{\infty} \left(1 - \frac{\sin \delta x}{\delta x}\right) dF_n(x) = \int_{-K}^{K} \left(1 - \frac{\sin \delta x}{\delta x}\right) dF_n(x) \\
&\quad + \int_{|x|>K} \left(1 - \frac{\sin \delta x}{\delta x}\right) dF_n(x)
\end{aligned}$$

tetszőleges $K > 0$ számra. Továbbá $1 - \frac{\sin \delta x}{\delta x} \ge 0$ minden x számra, és $1 - \frac{\sin \delta x}{\delta x} \ge \frac{1}{2}$, ha $x \ge \frac{2}{\delta}$. Innen kapjuk $K = \frac{2}{\delta}$ választással, hogy

$$\frac{\varepsilon}{2} \ge \frac{1}{2\delta}\int_{-\delta}^{\delta} \Re[1-\varphi_n(t)]\,dt \ge \int_{|x|>K} \frac{1}{2}\,dF_n(x),$$

azaz $F_n(-K) + (1 - F_n(K)) \le \varepsilon$, ha $n \ge n_0$. Nem nehéz belátni, hogy a K index esetleges növelésével elérhető, hogy ez az egyenlőtlenség a maradék véges sok $1 \le n \le n_0$ indexre is teljesüljön.