Az október 16-i gyakorlat témája
Rövid összefoglaló

1. Legyen ξ standard normális eloszlású valószínűségi változó, azaz legyen ξ sűrűségfüggvénye $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, és legyen t valós szám. Számoljuk ki az $e^{t\xi}$ valószínűségi változó $Ee^{t\xi}$ várható értékét.

   Megoldás: Ha ξ $f(x)$ sűrűségfüggvénnyel rendelkező valószínűségi változó $h(x)$ valós értékű (mérhető) függvény, akkor $Eh(\xi) = \int h(x)f(x)\,dx$. Esetünkben ez azt jelenti, hogy ha ξ standard normális eloszlású valószínűségi változó, akkor

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

   Innen

   $$Ee^{t\xi} = \int_{-\infty}^{\infty} \frac{e^{tx-x^2/2}}{\sqrt{2\pi}}\,dx = \int_{-\infty}^{\infty} \frac{e^{t^2/2-(t-x)^2/2}}{\sqrt{2\pi}}\,dx = e^{t^2/2} \int_{-\infty}^{\infty} \frac{e^{-x^2/2}}{\sqrt{2\pi}}\,dx = e^{t^2/2}.$$

Az első feladat megoldásához hasonlóan lehet tárgyalni a következő kérdést. Ennek fontos szerepe van a centrális határeloszlástétel bizonyításában.

2. Számítsuk ki egy standard normális eloszlású valószínűségi változó $E^{it\xi}$, $-\infty < t < \infty$, karakterisztikus függvényét. (Itt $i = \sqrt{-1}$.)

   Megoldás: Az első feladat megoldásához hasonlóan

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

   A feladat megoldásához elég belátni, hogy $\int_{-\infty+it}^{\infty+it} \frac{e^{-x^2/2}}{\sqrt{2\pi}}\,dx = 1$. Ez az állítás igaz, de a bizonyításhoz szükség van a komplex függvénytan néhány alapvető (nem triviális) eredményére.

   Az $f(z) = \frac{e^{-z^2/2}}{\sqrt{2\pi}}$ függvény analitikus az egész komplex számsíkon. Ezért a komplex számsík minden zárt görbéjére $\oint_\gamma f(z)\,dz = 0$. Válasszuk a következő γ_t görbéket. Vegyük először a $[-T,T]$ vízszintes szakaszt (pozitív irányban), majd menjünk tovább a $[T, T+it]$ függőleges szakaszon. Folytassuk az útunkat a $[T+it, -T+it]$ vízszintes szakaszon, majd zárjuk be a kört a $[-T+it, -T]$ szakaszon. Vegyük észre ezen kívül, hogy $z = x+iy$ esetben $|e^z| = e^x|e^{iy}| = e^x$, ahonnan $e^z \to 0$, ha $\operatorname{Re} z \to -\infty$. Ezt felhasználva kapjuk, hogy $\oint_{\gamma_t} f(z)\,dz = 0$, továbbá $t \to \infty$ határátmenet esetén a függőleges szakaszokon vett integrálokra $\int_{[\pm T, \pm T+it]} f(z)\,dz = 0$, ahonnan $\displaystyle\int_{-\infty+it}^{\infty+it} \frac{e^{-x^2/2}}{\sqrt{2\pi}}\,dx = \int_{-\infty}^{\infty} \frac{e^{-x^2/2}}{\sqrt{2\pi}}\,dx = 1.$

3. Adva egy ξ standard normális eloszlású valószínűségi változó számítsuk ki $8\xi^3$ és ξ^4 eloszlás és sűrűségfüggvényét.

a.) $8\xi^3$ eloszás és sűrűségfüggvénye: $P(8\xi^3 < x) = P\left(\xi < \frac{x^{1/3}}{2}\right) = \Phi\left(\frac{x^{1/3}}{2}\right)$, tehát $8\xi^3$ eloszlásfüggvénye $\Phi\left(\frac{x^{1/3}}{2}\right)$, sűrűségfüggvénye pedig ennek deriváltja, azaz $\frac{1}{6}x^{-2/3}\varphi\left(\frac{x^{1/3}}{2}\right) = \frac{x^{-2/3}}{6\sqrt{2\pi}}e^{-x^{2/3}/8}$.

b.) ξ^4 eloszlás és sűrűségfüggvénye: $P(\xi^4 < x) = 0$, ha $x \le 0$, és $P(\xi^4 < x) = P(-x^{1/4} < \xi < x^{1/4}) = P(\xi < x^{1/4}) - P(\xi < -x^{1/4}) = \Phi(x^{1/4}) - \Phi(-x^{1/4})$, ha $x > 0$. Másrészt, mivel a $\Phi(x)$ eloszlásfüggvény $\varphi(x)$ sűrűségfüggvénye páros függvény, ezért $\Phi(-x^{1/4}) = \displaystyle\int_{-\infty}^{-x^{1/4}} \varphi(u)\,du = \int_{x^{1/4}}^{\infty} \varphi(u)\,du = 1 - \int_{-\infty}^{x^{1/4}} \varphi(u)\,du = 1 - \Phi(x^{1/4})$. Ezért a ξ^4 valószínűségi változó eloszlásfüggvénye $P(\xi^4 < x) = 2\Phi(x^{1/4}) - 1$, ha $x \ge 0$, és nullával egyenlő, ha $x < 0$. Sűrűségfüggvénye pedig ennek deriváltja $\frac{1}{2}x^{-3/4}\varphi(x^{1/4})$, ha $x \ge 0$, és nulla, ha $x < 0$.

Házi feladat:

Legyen ξ egyenletes eloszlású valószínűségi változó a $[-\frac{1}{2}, \frac{1}{2}]$ intervallumban. Számítsuk ki a ξ^2 és ξ^3 valószínűségi változók eloszlás és sűrűségfüggvényét.

Házi feladat:

Legyen ξ és η két független λ paraméterű exponenciális eloszlású valószínűségi változó. Számoljuk ki $\xi - \eta$ sűrűségfüggvényét.

4. Egy államban, (nevezzük az egyszerűség kedvéért Floridának,) 5 000 000 választó választ két párt (hívjuk ezeket mondjuk republikánus és demokrata pártnak) jelöltje között. Tegyük fel, hogy a választók egymástól függetlenül $\frac{1}{2}$ valószínűséggel választják valamelyik párt jelöltjét. Mi annak a valószínűsége, hogy a két jelölt által összegyüjtött szavazatok különbsége nem haladja meg a háromszázat.

Az egyik jelöltre leadott szavazatok száma közelítőleg normális eloszlású $\frac{1}{2}\cdot 5\,000\,000$ várható értékkel és $\frac{1}{4}\cdot 5\,000\,000$ szórásnégyzettel. Annak a valószínűsége, hogy a két jelöltre adott szavazatok különbsége kisebb mint 300 megegyezik annak a valószínűségével, hogy az egyik jelöltre adott szavazatok száma a szavazatok várható értékétől kevesebb mint 150-nel tér el, ez pedig körülbelül, $\Phi\left(\frac{150}{\sqrt{\frac{1}{4}\cdot 5000000}}\right) -$

$$\Phi\left(-\frac{150}{\sqrt{\frac{1}{4}\cdot 5000000}}\right) = 2\Phi\left(\frac{6\sqrt{5}}{100}\right) - 1 \sim 2\Phi(0.124) - 1 \sim 0.1.$$

Megjegyzés: Érdemes megjegyezni, hogy igaz a centrális határeloszlás tétel következő lokális változata is a binomiális eloszlásra. Ha ξ binomiális eloszlású valószínűségi változó n és p paraméterekkel, akkor ξ várható értéke $mn = np$, szórásnégyzete $n\sigma^2 = np(1-p)$, és $P(\xi = k) = P\left(\frac{\xi - nm}{\sqrt{n}\sigma} = \frac{k-nm}{\sqrt{n}\sigma}\right) \sim \frac{1}{\sqrt{n}\sigma}\varphi\left(\frac{k-nm}{\sqrt{n}\sigma}\right) = \frac{1}{\sqrt{2\pi n}\sigma}\exp\left(\frac{(k-nm)^2}{2n\sigma^2}\right)$, ahol $\varphi(x) = \frac{1}{2\pi}e^{-x^2/2}$, a standard normális eloszlás sűrűségfüggvénye. Az előbb tekintett esetben $n = 5\,000\,000$, $p = \frac{1}{2}$, $-150 \le k \le 150$, valószínűségeket kell kiszámítani, és k-ra összegezni. Ebben az esetben a normális sűrűségfüggvényt a nulla közelében kell vennünk, és a kívánt valószínűésegre jó becsés $\frac{1}{\frac{1}{2}\sqrt{2\pi\cdot 5000000}}\cdot 300 \sim \frac{6}{10\sqrt{10\pi}} \sim \frac{1}{10}$.

2. Megjegyzés: Valójában az első feladatban tárgyalt modell némileg irreális. Általában vannak csoportok, melyeknek azonos a véleményük. Tegyük fel, hogy a különböző csoportokban levő emberek véleménye független, de az egyes csoportokban levő emberek véleménye megegyezik. Kérdés: Hogyan befolyásolja ez a tény annak valószínűségét, hogy rendkívül szoros választási eredmény szülessen? Be fogjuk látni, hogy e tény figyelembevétele csökkenti a szoros választási eredmény valószínűségét.

A fenti kérdés azzal függ össze, hogy ha különböző embereknek azonos a véleményük, akkor az egyik jelöltre leadott (véletlen, közel normális eloszlású) számának a szórása nő vagy csökken. A feladatot úgy is reprezentálhatjuk, hogy egyes csoportokból kijelölünk egy embert, az annyi szavazatot ad le, mint amennyi a csoport tagjainak a száma, a többi ember nem szavaz. Az, hogy ebben az esetben nem változik az egyik jelöltre leadott szavazatok várható értéke, de növekszik annak szórása, következik az alábbi feladat eredményéből.

5. Legyen $\xi_1, \dots, \xi_n$ független, egyforma eloszlású valószínűségi változók sorozata, melyeknek nem ismerjük a várható értékét. Lássuk be, hogy e valószínűségi változók tetszőleges súlyozott átlaga, azaz a $T_n(\xi_1, \dots, \xi_n) = \frac{\sum_{k=1}^n a_k\xi_k}{\sum_{k=1}^n a_k}$, $\sum_{k=1}^n a_k \ne 0$ a ξ_1 valószínűségi változó várható értékének torzítatlan becslése, melynek szórásnégyzete akkor a legkisebb, ha $a_1 = a_2 = \cdots = a_n$.

Megoldás: $E\frac{\sum_{k=1}^n a_k\xi_k}{\sum_{k=1}^n a_k} = \frac{\sum_{k=1}^n a_k E\xi_1}{\sum_{k=1}^n a_k} = E\xi_1$, és ez jelenti azt, hogy a várható érték torzítatlan becslését írtuk fel. Másrészt

$$\operatorname{Var}\frac{\sum\limits_{k=1}^{n}a_k\xi_k}{\sum\limits_{k=1}^{n}a_k}=\frac{\sum\limits_{k=1}^{n}a_k^2E\xi_1}{\left(\sum\limits_{k=1}^{n}a_k\right)}=\frac{\sum\limits_{k=1}^{n}a_k^2}{\left(\sum\limits_{k=1}^{n}a_k\right)^2}\operatorname{Var}\xi_1.$$

Ezért, ha $a_1 = a_2 = \cdots = a_n$, akkor $\operatorname{Var}\frac{\sum\limits_{k=1}^{n}a_k\xi_k}{\sum\limits_{k=1}^{n}a_k}=\frac{\operatorname{Var}\xi_1}{n}$. Másrészt, az úgynevezett Cauchy–Schwarz egyenlőtlenség alapján $\sum\limits_{k=1}^{n}a_k^2\le\frac{1}{n}\left(\sum\limits_{k=1}^{n}a_k\right)^2$, ahonnan

$$\operatorname{Var}\frac{\sum\limits_{k=1}^{n}a_k\xi_k}{\sum\limits_{k=1}^{n}a_k}\le\frac{\operatorname{Var}\xi_1}{n}$$

az általános esetben.

Eloszlásban való konvergencia fogalma és néhńy fontos tulajdonsága.

Az eloszlásban való konvergencia definiciója: *Valószínűségi változók ξ_n, $n = 1, 2, \dots$, sorozata akkor konvergál eloszlásban egy $F(u)$ eloszlásfüggvényhez vagy az ezen eloszlásfüggvény által meghatározott eloszláshoz, ha* $\lim\limits_{n\to\infty} P(\xi_n < u) = F(u)$ minden olyan u számra, ahol az $F(\cdot)$ eloszlásfüggvény függvény folytonos. *(Azt mondjuk, hogy a ξ_n, $n = 1, 2, \dots$, valószínűségi változók sorozata eloszlásban konvergál egy ξ valószínűségi változóhoz, ha ez a sorozat eloszlásban konvergál az $F(u) = P(\xi < u)$ eloszlásfüggvényhez.)*

Eloszlásfüggvények $F_n(u)$ sorozata akkor és csak akkor konvergál egy $F(u)$ eloszlásfüggvényhez, ha valószínűségi változók olyan ξ_n, $n = 1, 2, \dots$, sorozata, melyekre ξ_n eloszlása $F_n(u)$ eloszlásban konvergálnak az $F(u)$ eloszlásfüggvényhez. Más szavakkal ez azt jelenti, hogy $\lim\limits_{n\to\infty} F_n(u) = F(u)$ az $F(\cdot)$ függvény minden folytonossági pontjában.

Miért csak a határeloszlás folytonossági pontjaiban követeltük meg a konvergenciát az eloszlásban való konvergencia definiciójában?

Amikor eloszlásfüggvényeket vizsgálunk, akkor sokszor érdemes az eloszlásfüggvény helyett az általuk meghatározott valószínűségi mértéket tekinteni. Azaz, adva valamilyen $F(\cdot)$ eloszlásfüggvény, definiáljuk azt a $\mu = \mu_F$ valószínűségi mértéket a számegyenesen, melyre $\mu([a, b)) = F(b) - F(a)$. A mértékelmélet eredményeiből következik, hogy létezik ilyen valószínűségi mérték, és azt egyértelműen meghatározza az F eloszlásfüggvény. Ezt a valószínűségi mértéket szemléletesen úgy képzelhetjük el, mint egy tömegeloszlást a számegyenesen, melyre a teljes számegyenes tömege (súlya) 1, és a számegyenes egy részhalmazának a valószínűsége megegyezik a halmaz súlyával. Az, hogy F_n eloszlásfüggvények sorozata konvergál valamely F eloszláshoz valójában azt jelenti, hogy az általuk meghatározott μ_{F_n} tömegeloszlások konvergálnak a μ_F tömegeloszláshoz.

Tekintsük a következő egyszerű példát: Legyen a_n olyan monoton növekedő számsorozat, melyre $\lim\limits_{n\to\infty} a_n = a$ valamely $-\infty < a < \infty$ számra. Legyen μ_n az a valószínűségi mérték, mely az a_n pontba van koncentrálva, $n = 1, 2, \dots$, azaz $\mu_n(B) = 1$, ha $a_n \in B$, és $\mu_n(B) = 0$, ha $a_n \notin B$ a számegyenes tetszőleges mérhető részhalmazára. Hasonlóan, legyen μ az a pontba koncentrált mérték. Természetes azt várni, hogy ha alkalmasan definiáltuk az eloszlásban való konvergencia fogalmát, akkor a μ_n eloszlások konvergálnak a μ mértékhez. A μ_n mérték által meghatározott eloszlásfüggvény, azaz egy olyan ξ_n valószínűségi változónak az eloszlásfüggvénye, melyre $P(\xi_n = a_n) = 1$, az az $F_n(\cdot)$ eloszlásfüggvény, melyre $F_n(u) = 0$, ha $u \le a_n$, és $F_n(u) = 1$, ha $u > a_n$. Hasonlóan, a μ mérték által meghatározott eloszlásfüggvény az az $F(\cdot)$ függvény, melyre $F(u) = 0$, ha $u \le a$, és $F(u) = 1$, ha $u > a$. Könnyű ellenőrizni, hogy $\lim\limits_{n\to\infty} F_n(u) = F(u)$ minden $u \ne a$ számra ebben a példában, de ez a reláció az $u = a$ számra nem teljesül. Tehát a μ_n mértékek eloszlásban konvergálnak ebben a példában, de ehhez szükséges volt az eloszlásban való konvergencia definiciójában azt a megszorítást tenni, hogy csak határeloszlásfüggvény folytonossági pontjaiban követeljük meg a konvergenciát.

A $\Phi(x) = \displaystyle\int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\,du$ normális eloszlásfüggvény minden pontban folytonos, ezért a centrális határeloszlástétel alkalmazásában nincs jelentősége annak, hogy az eloszlásban való konvergencia csak az eloszlásfüggvény folytonossági pontjaiban követelünk meg konvergenciát. Viszont olyan határeloszlástételben, melyben a határeloszlásnak van szakadási pontja (például a Poisson eloszlás vagy bármely diszkrét eloszlás ilyen) ez a megszorítás lényeges. Kimondunk bizonyítás nélkül egy eredményt, mely ekvivalens feltételt ad arra, hogy eloszlásfüggvények egy sorozata eloszlásban konvergáljon egy határeloszláshoz. Ennek az eredménynek fontos szerepe van a karakterisztikus függvény módszer alkalmazásában is.

Tétel. *$F_n(u)$, $n = 1, 2, \dots$ eloszlásfüggvények sorozata akkor és csak akkor konvergál eloszlásban egy $F(u)$ eloszlásfüggvényhez, ha minden a számegyenesen értelmezett folytonos, korlátos $g(u)$ függvényre*

$$\lim_{n\to\infty} \int g(u)\,dF_n(u) = \int g(u)\,dF(u).$$

Eloszlásban való konvergencia és határeloszlástételek.

Vegyük észre, hogy mivel az $f_t(x) = e^{itx}$, $-\infty < t < \infty$, függvények folytonosak és korlátosak, ezért az eloszlásfüggvényeknek a fenti Tétel-ben kimondott jellemzése alapján eloszlások konvergenciájából következik, hogy $\lim\limits_{n\to\infty} \int e^{itx} F_n(\,dx) = \int e^{itx} F(\,dx)$, minden $-\infty < t < \infty$ számra, ha az F_n eloszlásfüggvények eloszlásban konvergálnak az F eloszlásfüggvényhez. Felmerül a kérdés: Igaz-e ennek az állításnak a megfordítása is, azaz, ha tudjuk, hogy a $\varphi_n(t) = \varepsilon^{itx} F_n(dx) \to \varphi(t) = \int e^{itx} F(\,dx)$, ha $n \to \infty$ reláció teljesül egy $F_n(\cdot)$, $n = 1, 2, \dots$, eloszlásfüggvénysorozatra valamint egy F eloszlásfüggvényre, akkor következik-e ebből az, hogy az F_n eloszlásfüggvények eloszlásban konvergálnak az F eloszlásfüggvényhez? Az előző formulában bevezetett $\varphi(t) = \int e^{itx} F(\,dx)$ függvényt nevezzük az F eloszlás karakterisztikus függvényének.

A valószínűségszámítás egyik alapvető fontosságú eredménye szerint a válasz erre a kérdésre igenlő. Jegyezzük meg, hogy ezen állításnak a bizonyításához szükséges egy olyan jellegű eredmény, mely azt mondja ki, hogy az $e(t) = e^{itx}$ trigonometrikus függvények az összes folytonos és korlátos függvényekből álló halmaznak elég nagy részhalmaza. Ugyanis, ha $\lim\limits_{n\to\infty} \int e^{itx} F_n(\,dx) = \int e^{itx} F(\,dx)$ minden $-\infty < t < \infty$ számra, akkor teljesül a $\lim\limits_{n\to\infty} \displaystyle\int \sum_{k=1}^{p} c_k e^{it_k x} F_n(\,dx) = \int \sum_{k=1}^{p} c_k e^{t_k x} F(\,dx)$ reláció minden véges lineáris kombinációra is, sőt bízonyos limeszeléssel eljutunk további olyan $g(\cdot)$ függvényekhez, melyekre $\lim\limits_{n\to\infty} \int g(x) F_n(\,dx) = \int g(x) F(\,dx)$. De eljutunk-e így minden folytonos és korlátos függvényhez. A valószínűségszámítás egyik mély eredménye szerint igen.

Eloszlások konvergenciájáról szóló alaptétel. *Legyen ξ_n, $n = 1, 2, \dots$, F_n eloszlású valószínűségi változók sorozata $\varphi_n(t) = Ee^{it\xi_n} = \int e^{itx} F_n(dx)$ karakterisztikus függvénnyel. Ha a $\varphi_n(t)$ függvények minden $-\infty < t < \infty$ számra konvergálnak egy $\varphi(t)$ függvényhez, mely az origóban folytonos függvény, akkor ez a $\varphi(t)$ limeszfüggvény is karakterisztikus függvény, azaz létezik olyan (egyértelműen meghatározott) $F(x)$ eloszlásfüggvény, melyre $\varphi(t) = \int e^{itx} F(\,dx)$. Továbbá, az F_n eloszlásfüggvények eloszlásban konvergálnak ehhez az F eloszlásfüggvényhez.*

Megfordítva, ha F_n eloszlások, $n = 1, 2, \dots$, sorozata eloszlásban konvergál egy F eloszláshoz, akkor minden $-\infty < t < \infty$ számra a $\varphi_n(t) = \int e^{itx} F_n(dx)$ karakterisztikus függvények konvergálnak a $\varphi(t) = \int e^{itx} F(dx)$ karakterisztikus függvényhez, ha $t \to \infty$. Továbbá ez a konvergencia egyenletes minden véges intervallumban.

Talán érdemes lett volna ennek az eredménynek a megfogalmazása előtt kimondani a következő tételt.

Tétel. *Egy F eloszlásfüggvényt meghatároz annak $\varphi(t) = \int e^{itx} F(\,dx)$ karakterisztikus függvénye.*