A szeptember 18-i gyakorlat témája
Rövid összefoglaló

Tekintsük a következő problémát:

Legyen egy urnában 70 piros és 30 fehér golyó. Húzzuk ki a golyókat visszatevés nélkül.

a.) Mi a valószínűsége annak, hogy az első kihúzott golyó színe piros, a másodiké pedig fehér?

b.) Mi a valószínűsége annak, hogy a 13. kihúzott golyó színe piros, a 29. kihúzott golyó színe pedig fehér?

A feladat b.) része az érdekes. Mutassuk meg, hogy az a.) és b.) kérdésre ugyanaz a válasz. Ezen állítás bizonyítása és megértése érdekében megfogalmaztuk a feladatot formálisan egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező bevezetésével.

A következő modellt definiáltuk: Tekintsük az összes $\omega = (F, P, F, F, \dots)$ száz hosszúságú 30 F (fehér) és 70 P (piros) jelből álló sorozatot. Legyen Ω az összes ilyen ω sorozatól álló halmaz, $\mathcal{A}$ a (mérhető) halmazok rendszere (σ-algebrája) Ω összes részhalmaza, azaz minden lehetséges fent definiált ω (elemi eseményekből) álló halmaz. Legyen minden ω elemi esemény valószínűsége $\frac{1}{\binom{100}{30}}$. Egy A halmaz valószínűsége pedig az A halmaz által tartalmazott elemi események valószínűségének összege. A lényegs észrevétel az volt, hogy minden olyan húzássorozatnak a valószínűsége, amelyik 30 fehér és 70 piros golyót tartalmaz megegyezik. Ebben a modellben az a.) feladat valószínűsége azon A esemény valószínűsége, mely azokat az ω sorozatokat tartalmazza, melyek első eleme P második eleme pedig F. A b.) feladatban annak a B eseménynek a valószínűségét vizsgáljuk, melyek olyan ω sorozatokat tekint, melyek 13. jele P 29. jele pedig F. Beláttuk két módon is, hogy $P(A) = P(B)$. Ehhez azt látjuk be, hogy azon sorozatok melyekben az első helyen piros és a második helyen fehér gplyó van kölcsönösen egyértelműen megfeleltethetőek azon sorozatoknak, melyeknek a 13. helyén piros és a 29. helyén fehér golyó áll. Ilyen megfeleltetést biztosít az, ha a 1. és 13. valamint a 2. és 29. golyót kicseréljük. Egy másik szintén megtárgyalt bizonyítása annak, hogy ugyanannyi az első helyen piros és a második helyen fehér golyót tartalmazó sorozat van, mint ahány a 13. helyen piros és 29. helyen fehér golyó áll, abból áll, hogy kiszámoljuk mind a kétfajta 100 hosszúságú sorozat számát. Azt kell észrevenni, hogy ugyanolyan módon lehet kiszámolni ezt a két lehetséges számot. Ugyanis az ilyen sorozatok számát úgy számolhatjuk ki, hogy tekintjük az összes olyan sorozatok számát, melyekben két helyen elő van írva a golyók színe, és ezután azt kell kiszámolni, hogy hány olyan sorozat van, amelyikben a maradék 68 helyen 29 fehér és 69 piros golyó van. Annak, hogy melyik két helyen van a golyók színe előírva nincs jelentősége.

Valójában a fenti tárgyalás nem teljesen kielégítő. Látnunk kell ugyanis, hogy nincs jelentősége annak, hogy a minket érdeklő valószínűségeket milyen a feltételeket kielégítő modellben tekintjük. Ezért érdemes a fenti gondolatmenetet úgy módosítani, hogy az érvelés ne függjön attól, hogy milyen modellt tekintünk.

Legyen adva egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező, mely tartalmazza az összes olyan eseményt, hogy egy húzássorozat eredménye valamilyen $(F, P, F, F, \dots)$ száz hosszúságú 30 F (fehér) és 70 P (piros) jelből álló sorozat. (Nem tesszük fel, hogy ezek az események elemi események.) Minden ilyen esemény valószínűsége legyen $\frac{1}{\binom{100}{30}}$. Vegyük észre, hogy ilyen módon a valószínűségi mező egy particióját kapjuk. (Diszjunkt eseményeket definiáltunk, melyek összvalószínűsége 1.) Definiáljuk a következő ξ_j, $1 \le j \le 100$, valószínűségi változókat: $\xi_j = 1$, ha a j-ik húzás eredménye piros, $\xi_j = 0$, ha a j-ik húzás eredménye fehér. (Azaz, ha az előbb definiált partició olyan sorozatot tartalmaz, melynek j-ik jele P illetve F.) Ekkor a minket érdeklő kérdés a $P(\xi_1 = 1, \xi_2 = 0)$ (a.) feladat), illetve a $P(\xi_{13} = 1, \xi_{29} = 0)$ valószínűség (b.) feladat) kiszámítása.

Be lehet látni, hogy a fenti jelölésekkel

$$P(\xi_1 = \varepsilon_1, \xi_2 = \varepsilon_2, \dots, \xi_{100} = \varepsilon_{100}) = P(\xi_{\pi(1)} = \varepsilon_1, \xi_{\pi(2)} = \varepsilon_2, \dots, \xi_{\pi(100)} = \varepsilon_{100}),$$

ahol $(\pi(1), \dots \pi(100))$ az $(1, \dots, 100)$ számok tetszőleges permutációja, ε_j, $1 \le j \le 100$ vagy nullla vagy egy. Ez az azonosság lehetővé teszi annak bizonyítását, hogy az a) és b) feladatban szereplő valószínűségek megegyeznek.

Házi feladat

Ha a $\xi_1, \xi_2, \dots, \xi_n$ egész értékű valószínűségi változók teljesítik a

$$P(\xi_1 = k_1, \xi_2 = k_2, \dots, \xi_n = k_n) = P(\xi_{\pi(1)} = k_1, \xi_{\pi(2)} = k_2, \dots, \xi_{\pi(n)} = k_n),$$

azonosságot az $1, 2, \dots, n$ számok tetszőleges $(\pi(1), \dots \pi(n))$ permutációjára és minden $k_1, \dots, k_n$ egész számra, akkor

$$P(\xi_i = k_1) = P(\xi_1 = k_1), \quad P(\xi_i = k_1, \xi_j = k_2) = P(\xi_1 = k_1, \xi_2 = k_2)$$

mainden $1 \le i, j \le n$ és k_1 és k_2 egész számra.

További a gyakorlaton tárgyalt feladatok:

1.) Feldobunk egy szabályos pénzdarabot 100-szor egymás után. Számítsuk ki a fejdobások számának várható értékét és szórásnégyzetét.

2.) Feldobunk egy szabályos dobókockát 100-szor egymás után. Tekintsük a dobások eredményeinek összegét. Számítsuk ki ennek az összegnek

a.) Várható értékét,

b.) Szórásnégyzetét.

Idézzük fel először a várható érték és szóráanégyzet definicióját és néhány fontos velük kapcsolatos eredményt.

Várható érték és szórásnégyzet definiciója. *Legyen ξ diszkrét érékű valószínűségi változó, mely megszámlálható sok $x_1, x_2, \dots$ értéket vesz fel, $P(\xi = x_k) = p_k$, $k =$*

$1, 2, \dots,$ *akkor a* ξ *valószínűségi változó várható értéke* $E\xi = \sum_{k=1}^{\infty} p_k x_k$, *szórásnégyzete pedig* $\operatorname{Var}\xi = E(\xi - E\xi)^2 = E\xi^2 - (E\xi)^2$.

Tétel. $E(a\xi + b) = aE\xi + b$, $\operatorname{Var}(a\xi + b) = a^2 \operatorname{Var}\xi$ *minden valós* a *és* b *számra.*

Tétel. *Ha* ξ *diszkrét értékű valószínűségi változó,* $P(\xi = x_k) = p_k$, $k = 1, 2, \dots,$ $f(x)$ *valamilyen függvény, akkor* $Ef(\xi) = \sum_{k=1}^{\infty} p_k f(x_k)$.

Tétel. *Ha* $\xi_1, \dots, \xi_n$ *valószínűségi változók ugyanazon a valószínűségi mezőn, akkor* $E(\xi_1 + \dots + \xi_n) = E\xi_1 + \dots + E\xi_n$.

Megjegyzés; Ez az eredmény tetszőleges valószínűségi változókra igaz. Nem követeltük meg például azt, hogy az összeadandók függetlenek legyenek.

Tétel. *Ha* $\xi_1, \dots, \xi_n$ *független valószínűségi változók, akkor* $\operatorname{Var}(\xi_1 + \dots + \xi_n) = \operatorname{Var}\xi_1 + \dots + \operatorname{Var}\xi_n$.

Az 1. és 2. feladat jobb megértése érdekében definiáljunk egy valószínűségi mezőt és rajta valószínűségi változókat, melyek segítségével ezt a feladatot precizen meg lehet fogalmazni.

1. *A feladat megoldása:* A fejdobások száma 0 és í00 között van. Annak valószínűsége, hogy k fejdobás következik be, $0 \le k \le 100$, $p_k = \binom{100}{k} 2^{-100}$. Ezért a dobások számának várható értéke $E\xi = \sum_{k=0}^{100} k \binom{100}{k} 2^{-100}$, ahol ξ jelőli a fejdobások számát megadó valószínűségi változót. Továbbá $E\xi^2 = \sum_{k=0}^{100} k^2 \binom{100}{k} 2^{-100}$, a szórásnégyzet pedig $\operatorname{Var}\xi = E\xi^2 - (E\xi)^2$. Ezeket az összegeket közvetlenül kiszámíthatjuk. Valóban, $k\binom{100}{k} = 100\binom{99}{k-1}$, $E\xi = \sum_{k=0}^{100} k\binom{100}{k} 2^{-100} = 50 \sum_{k=0}^{99} k \binom{99}{k} 2^{-99} = 50\left(\frac{1}{2} + \frac{1}{2}\right)^{99} = 50$. Továbbá, $k^2\binom{100}{k} = [k(k-1)+k]\binom{100}{k} = 100 \cdot 99 \cdot \binom{98}{k-2} + 100 \cdot \binom{100}{99}$, $E\xi^2 = \frac{1}{4} \cdot 100 \cdot 99 \cdot \sum_{k=0}^{98} \binom{98}{k} + \frac{1}{2} \cdot 100 \cdot \sum_{k=0}^{99} \binom{99}{k} = 25 \cdot 99 + 50$, $\operatorname{Var}\xi = E\xi^2 - (E\xi^2) = 25$.

   Valójában a vizsgált várható értéket és szórásnégyzetet egyszerűbben is kiszámíthatjuk. Vezessük be a ξ_j valószínűségi változókat, $\xi_j = 1$, ha a j-ik dobás eredménye fej $\xi_j = 0$, ha a j-ik dobás eredménye írás, $1 \le j \le 100$. Ekkor a fejdobások száma $\xi = \sum_{j=1}^{100} \xi_j$, $E\xi = \sum_{j=1}^{100} E\xi_j$, $\operatorname{Var}\xi = \sum_{j=1}^{100} \operatorname{Var}\xi_j$. Mivel $E\xi_j = \frac{1}{2}$, $E\xi_j^2 = \frac{1}{2}$, $\operatorname{Var}\xi_j = \frac{1}{4}$, $1 \le j \le 100$, ezért $E\xi = \frac{1}{2} \cdot 100 = 50$, $\operatorname{Var}\xi = \frac{1}{4} \cdot 100 = 25$.

A második feladat hasonlóan tárgyalható. Ekkor annak valószínűségét, hogy a dobások üsszegének értéke egy adott szám rendkívül fáradságos lenne kiszámolni. Viszont a második módszer egyszerűen alkalmazható ebben az esetben is. Legyenek $\eta_1,\dots,\eta_{100}$ *független* valószínűségi változók, melyekre $P(\eta_j = k) = \frac{1}{6}$, $1 \le j \le 100$, $1 \le k \le 6$. Ekkor a kiszámítandó várható érték és szórásnégyzet $E\sum\limits_{j=1}^{100}\eta_j = \sum\limits_{j=1}^{100}E\eta_j$, $\operatorname{Var}\sum\limits_{j=1}^{100}\eta_j = \sum\limits_{j=1}^{100}\operatorname{Var}\eta_j$. (A második reláció felhasználja a tekintett valószínűségi változók függetlenségét.) $E\eta_j = \frac{1}{6}(1+2+3+4+5+6) = 3.5$, $E\eta_j^2 = \frac{1}{6}(1+4+9+16+25+36) = \frac{91}{6}$, $\operatorname{Var}\eta_j = \frac{35}{12}$, $E\xi = 350$, $\operatorname{Var}\xi = \frac{3500}{12}$.

3.) Legyen egy urnában 40 piros és 60 fehér golyó. Húzzuk ki a golyókat visszatevés nélkül. Számítsuk ki az első 10 húzásban kihízott piros golyók számának

a.) Várható értékét,

b.) Szórásnégyzetét.

Az első két feladathoz hasonlóan tárgyalható a harmadik feladat. Ekkor megfelelő modellben a vizsgált feladat az $E\sum\limits_{j=1}^{20}\xi_j$ és $\operatorname{Var}\sum\limits_{j=1}^{20}\xi_j$ kiszámítása, ahol $\xi_j = 1$, ha a j-ik húzás eredménye piros, $\xi_j = 0$, ha a j-ik húzás eredménye fehér. Ekkor $E\sum\limits_{j=1}^{20}\xi_j = \sum\limits_{j=1}^{20}E\xi_j$. A szórásnégyzet kiszámítása kissé bonyolultabb, mert a tekintett valószínűségi változók nem függetlenek. Ezért meg kell értenünk, hogyan kell kiszámítani valószínűségi változók összegének a szórásnégyzetét az általános esetben.

A $\xi = \sum\limits_{j=1}^{20}\xi_j$ összeg szórásnégyzetét kell kiszámolnunk, ahol $\xi_j = 1$, ha a j-ik húzás eredménye piros, $\xi_j = 0$ ha a j-ik húzás eredménye fehér, $j = 1,2,\dots,20$. Ennek kiszámítása bonyolultabb mint az előző feladatokban, mert a most tekintett ξ_j valószínűségi változók nem függetlenek. Ezért meg kell értenünk, hogyan kell kiszámítani valószínűségi változók összegének a szórásnégyzetét az általános esetben.

Tétel. *Ha $\xi_1,\dots,\xi_n$ valószínűségi változók ugyanazon a valószínűségi mezőn, akkor*

$$\begin{aligned}\operatorname{Var}\left(\sum_{j=1}^{n}\xi_j\right) &= \sum_{j=1}^{n}\operatorname{Var}\xi_j + 2\sum_{1\le j<k<n}(E\xi_j\xi_k - E\xi_jE\xi_k).\\ &= \sum_{j=1}^{n}\operatorname{Var}\xi_j + \sum_{1\le j,k\le n,\ j\ne k}(E\xi_j\xi_k - E\xi_jE\xi_k).\end{aligned}$$

Megjegyzés: Ha ξ és η két valószínűségi változó, akkor ezek kovarianciája $\operatorname{Cov}(\xi,\eta) = E[(\xi - E\xi)(\eta - E\eta)] = E\xi\eta - E\xi E\eta$. A kovariancia egy természetes mérése annak, hogy két valószínűségi változó milyen mértékben függ össze. Ez a mennyiség szerepelt az összeg szórásnégyzetének kifejezésében. Ha ξ és η független valószínűségi változók, akkor $\mathrm{Cov}\,(\xi,\eta)=0$.

A Tétel bizonyítása.

$$\mathrm{Var}\left(\sum_{j=1}^n \xi_j\right) = E\left(\sum_{j=1}^n \xi_j\right)^2 - \left(E\sum_{j=1}^n \xi_j\right)^2 = E\left(\sum_{j=1}^m\sum_{k=1}^n \xi_j\xi_k\right) - \left(\sum_{j=1}^m\sum_{k=1}^n E\xi_j E\xi_k\right)$$
$$= \sum_{j=1}^n \left(E\xi_j^2 - (E\xi_j)^2\right) + 2\sum_{1\le j<k<n} (E\xi_j\xi_k - E\xi_j E\xi_k).$$

és $E\xi_j^2 - (E\xi_j)^2 = \mathrm{Var}\,\xi_j$.

Alkalmazva ezt a tételt és az előző gyakorlat eredményeit kapjuk, hogy $E\xi_j = E\xi_1 = \frac{2}{5}$, $\mathrm{Var}\,\xi_j = \mathrm{Var}\,\xi_1 = \frac{2}{5} - \frac{4}{25} = \frac{6}{25}$, $E\xi_j\xi_k - E\xi_j E\xi_k = E\xi_1\xi_2 - E\xi_1 E\xi_2 = \frac{2}{5}\cdot\frac{39}{99} - \frac{4}{25} = -\frac{6}{495}$, ha $j\neq k$, és $E\xi_j\xi_k - E\xi_j E\xi_k = E\xi_1\xi_2 - E\xi_1 E\xi_2$ minden $1\le j<k\le 20$ számra. Innen a 10 dobásban kihúzott piros golyók számának várható értéke $10\cdot\frac{2}{5}=4$ és szórásnégyzete $10\cdot\frac{6}{25} - 90\cdot\frac{6}{2475} = \frac{12}{5}\left(1-\frac{1}{11}\right) = \frac{24}{11}$.

Házi feladat.

Egy urnában húsz piros és harminc fehér golyó van. Kihúzunk egymás után 10 golyót úgy, hogy amikor kihúzunk egy golyót akkor azt visszadobjuk és vele együtt bedobunk az urnába egy vele azonos színű golyót. Mi a kihúzott piros golyók számának a várható értéke és szórásnégyzete?

Segítség: Lássuk be, hogy annak valószínűsége, hogy a j-ik húzásban piros golyót húzunk megegyezik annak valószínűségével, hogy az első húzásban piros golyót húzunk. Annak valószínűsége, hogy az i-ik és j-ik húzásban piros golyót húzunk, $i\neq j$, megegyezik annak valószínűségével, hogy az első és második húzásban piros golyót húzunk.