A gyakorlaton tárgyalt feladatok elméleti hátteréről.

A gyakorlaton tárgyalt feladatok vizsgálatában nagyon fontos a függetlenség fogalmának, (mind az események mind a valószínűségi változók függetlenségének) az ismerete, illetve az e fogalmakhoz kapcsolódó eredmények használata. Ugyancsak fontos a várható érték és szórásnégyzet fogalmának az ismerete, illetve az, hogy az e mennyiségek kiszámításához szükséges eredményeket tudjuk jól használni. Valószínűségi változók függetlenéségét valamint ezek várható értékét és szórásnégyzetét egyszerűbb először diszkrét eloszlású valószínűségi változók esetén tárgyalni és csak azután térni rá az általános esetre. (Ezt a tárgyalási módot követi Fazekas István jegyzete is.) Egy másik, rövidebb tárgyalási mód, rögtön az általános eset ismertetése. Ezzel sok ismétlést el lehet kerülni. Ezt a hozzáállást követi az előadás is. Ugyanakkor fontos megérteni e tárgyalási mód esetén is at, hogy mit mondanak ki az általános eredmények abban a speciális esetben, ha diszkrét eloszlású valószínűségi változókat tekintünk. Az alábbiakban ismertetem a legfontosabb tudnivalókat.

Két esemény függetlenségének a definiciója. *Azt mondjuk, hogy egy A és B esemény független, ha $P(A \cap B) = P(A)P(B)$.*

Ez a definició azonban önmagában nem kielégítő számunkra. Beszélni akarunk több esemény függetlenségéről is. Ezért a következő definiciót is bevezetjük.

Több esemény függetlenségének definiciója: *Az $A_1, \dots, A_n$ események akkor (teljesen) függetlenek, ha az $\{1, \dots, n\}$ indexhalmaz minden $\{j_1, \dots, j_k\} \subset \{1, \dots, n\}$ részhalmazára*

$$P(A_{j_1} \cap \cdots \cap A_{j_k}) = P(A_{j_1}) \cdots P(A_{j_k}).$$

Események végtelen $A_1, A_2, \dots$ sorozata akkor és csak akkor független, ha tetszőleges pozitív n egész számra az $A_1, \dots, A_n$ események függetlenek.

Speciálisan $n = 3$ esetben ez a definició a következőt jelenti: Az A, B és C események akkor függetlenek, ha a következő azonosságok *mindegyike* teljesül:

$$\begin{aligned} P(A \cap B) &= P(A)P(B) \\ P(A \cap C) &= P(A)P(C) \\ P(B \cap C) &= P(B)P(C) \\ P(A \cap B \cap C) &= P(A)P(B)P(C). \end{aligned}$$

Bizonyos alkalmazásokban fontos a következő eredmény.

Lemma. *Ha $A_1, \dots, A_n$ független események, és bevezetjük az $A_j^1 = A_j$ és $A_j^{-1} = \Omega \setminus A_j$ jelöléseket, akkor tetszőleges $\varepsilon_j = \pm 1$, $1 \le j \le n$, sorozatra az $A_1^{\varepsilon_1}, \dots, A_n^{\varepsilon_n}$ események függetlenek.*

Vezessük be először diszkrét eloszlású valószínűségi változók fogalmát, majd azok eloszlását, függetlenségét.

Először felidézek néhány fontos eredményt (diszkrét eloszlású) valószínűségi változók illetve ilyen valószínűségi változók összegének a várható értékéről és szórásnégyzetéről.

Valószínűségi változó fogalma: *Legyen adva egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező. Egy ezen a téren értelmezett valós értékű (mérhető) $\xi(\omega)$ függvényt valószínűségi változónak fogunk nevezni. Azaz ξ az $\Omega \to R^1$ leképezés. Az a megkötés, hogy a függvény legyen mérhető azt jelenti, hogy minden x valós számra az $\{\omega\colon \xi(\omega) < x\} \in \mathcal{A}$ feltétel teljesül.*

Időnként hasznos nemcsak valós értékű valószínűségi változóról beszélni, hanem kissé általánosabban olyan valószínűségi változókat definiálni, amelyek értékeiket egy általános téren veszik fel. (Például bizonyos esetekben érdemes olyan valószínűségi változókat tekinteni, amelyek értéke komplex szám, vagy nem egyetlen szám, hanem egy szám-n-es. Az ilyen valószínűségi változókat vektor értékűnek szokták hívni.) Annak érdekében, hogy ezt megtehessük tekintsünk valamilyen X halmazt, és e halmaz kitüntetett részhalmazainak $\mathcal{X}$ osztályát, amelyek σ-algebrát alkotnak. Egy az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn definiált X térbeli értékeket felvevő (mérhető) függvényt X-tér értékű valószínűségi változónak nevezünk. Az, hogy ez a függvény mérhető azt jelenti, hogy minden $Y \in \mathcal{X}$ halmazra az $\{\omega\colon \xi(\omega) \in Y\} \in \mathcal{A}$ feltétel teljesül.

Diszkrét eloszlású valószínűségi változók és azok eloszlása. *Egy ξ valószínűségi változót egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn diszkrétnek vagy diszkrét eloszlásúnak hívunk, ha megadható egy véges vagy megszámlálhatóan végtelen számosságú $x_1, x_2, \cdots,$ halmaz úgy, hogy az $\{\omega, \xi(\omega) = x_n\}$, $n = 1, 2, \dots,$ halmazok teljes eseményrendszert alkotnak. Egy diszkrét eloszlású valószínűségi változó eloszlását meghatározzák azok az $x_1, x_2, \dots,$ értékek amelyeket az felvesz és a $p_n = P(\xi = x_n)$ valószínűségek. (Jegyezzük meg, hogy $\sum\limits_n P(\xi = x_n) = 1$.)*

Legyenek $\xi_1, \dots, \xi_k$ diszkrét eloszlású valószínűségi változók ugyanazon az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, amelyek valamilyen $x_1, x_2, \dots$ értékeket vesznek fel. Ezek együttes eloszlását meghatározzák a $P(\xi_1 = x_{j_1}, \xi_2 = x_{j_2}, \dots, \xi_k = x_{j_k})$, $j_s = 1, 2, \dots,$ $1 \le s \le k$, valószínűségek.

Diszkrét eloszlású valószínűségi változók függetlensége. *Legyenek $\xi_1, \dots, \xi_k$ diszkrét eloszlású valószínűségi változók ugyanazon az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, amelyek valamilyen $x_1, x_2, \dots$ értékeket vesznek fel. Azt mondjuk, hogy ezek az ξ_1, ξ_2, ... , ξ_k valószínűségi változók függetlenek, ha*

$$P\left(\xi_1 = x_{j_1}, \xi_2 = x_{j_2}, \dots, \xi_k = x_{j_k}\right) = P\left(\xi_1 = x_{j_1}\right) P\left(\xi_2 = x_{j_2}\right) \cdots P\left(\xi_k = x_{j_k}\right)$$

minden $j_s = 1, 2, \dots,$ $1 \le s \le k$, indexre.

Igazak a következő eredmények.

Lemma: *Legyenek adva $\xi_1, \dots, \xi_k$ független, diszkrét eloszlású valószínűségi változók egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, amelyek valamilyen $x_1, x_2, \dots$ értékeket vesznek fel. Legyenek $A_1 \subset \{x_1, x_2, \dots\}, A_2 \subset \{x_1, x_2, \dots\}, \dots, A_k \subset \{x_1, x_2, \dots\}$ tetszőleges halmazok. Ekkor*

$$P\left(\xi_1 \in A_1, \xi_2 \in A_2, \dots, \xi_k \in A_k\right) = P\left(\xi_1 \in A_1\right) P\left(\xi_2 \in A_2\right) \cdots P\left(\xi_k \in A_k\right).$$

Speciálisan, ekkor tetszőleges $\{j_1, \dots, j_s\} \subset \{1, \dots, k\}$ indexhalmazra a $\xi_{j_1}, \dots, \xi_{j_s}$ valószínűségi változók függetlenek.

Tétel: *Legyenek adva $\xi_1, \dots, \xi_k$ független, diszkrét eloszlású valószínűségi változók egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, és legyenek $h(x_1), \dots, h(x_k)$ tetszőleges függvények, amelyek a $\xi_1, \dots, \xi_k$ valószínűségi változók értékkészletén vannak definiálva, ezért lehet az $\eta_1 = g_1(\xi_1), \dots, \eta_k = g_k(\xi_k)$ valószínűségi változókról beszélni. Ekkor $\eta_1, \dots, \eta_k$ független valószínűségi változók.*

Halmaz indikátorfüggvényének a definiciója. *Legyen adva egy $A \in \mathcal{A}$ esemény egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Az A halmaz indikátorfüggvényén azt a $\chi_A(\omega)$ valószínűségi változót értjük, amelyre $\chi_A(\omega) = 1$, ha $\omega \in A$, és $\chi_A(\omega) = 0$, ha $\omega \notin A$.*

Ezzel a jelöléssel a következő kapcsolatot létesíthetjük események és (diszkrét eloszlású) valószínűségi változók függetlensége között.

Lemma. *Legyenek $A_1, \dots, A_k$ események egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Az $A_1, \dots, A_k$ események akkor és csak akkor függetlenek, ha azok $\chi_{A_1}(\omega), \dots, \chi_{A_k}(\omega)$ indikátorfüggvényei függetlenek.*

Rátérek a várható érték és szórásnégyzet tárgyalására.

Diszkrét eloszlású valószínűségi változó várható értéke. *Legyen ξ diszkrét eloszlású valószínűségi változó, amely $x_1, x_2, \dots$ értékeket vesz fel $p_k = P(\xi = x_k)$ valószínűséggel, $k = 1, 2, \dots$, $\sum\limits_{k=1}^{\infty} P(\xi = x_k) = 1$. A ξ valószínűségi változó várható értéke az*

$$E\xi = \sum_{k=1}^{\infty} x_k P(\xi = x_k)$$

összeg, feltéve hogy ez az összeg abszolut konvergens. Ha ez az összeg nem abszolut konvergens, akkor nem definiáljuk az $E\xi$ várható értéket.

Tétel. *Legyenek ξ_1, ξ_2 (diszkrét eloszlású) valószínűségi változók, amelyeknek létezik várható értékük. Ekkor a $\xi_1 + \xi_2$ összegnek is létezik várható értéke, és*

$$E(\xi_1 + \xi_2) = E\xi_1 + E\xi_2.$$

Következmény. *Legyenek adva $\xi_1, \xi_2, \dots, \xi_k$ (diszkrét eloszlású) valószínűségi változók, amelyeknek létezik várható értékük, és legyenek $c_1, \dots, c_k$ valós számok. Ekkor a $c_1\xi_1 + c_2\xi_2 + \dots + c_k\xi_k$ valószínűségi változónak is létezik várható értéke, és*

$$E\left(c_1\xi_1 + c_2\xi_2 + \dots + c_k\xi_k\right) = c_1 E\xi_1 + c_2 E\xi_2 + \dots + c_k E\xi_k.$$

Megjegyzés: A fenti tétel, illetve annak következménye érvényes általános, nem feltétlenül diszkrét eloszlású valószínűségi változók esetében is.

Fontos észrevétel: A fenti tételbem, illetve annak következményében nem tettük fel, hogy az összegben szereplő valószínűségi változók függetlenek. Ez azt jelenti, hogy valószínűségi változók összegének a várható értékét ki tudjuk számítani akkor is, ha csak az egyes összeadandók várható értékét tudjuk kiszámolni. Az, hogy az egyes összeadandók függetlenek-e vagy erősen függnek-e egymástól, nem befolyásolja összegük várható értékének a nagyságát.

Tétel. *Legyen ξ diszkrét eloszlású valószínűségi változó, amely bizonyos $x_1, x_2, \ldots$ értéket vesz fel és $g(x)$ valós függvény. Ekkor*

$$Eg(\xi) = \sum_{j=1}^{\infty} g(x_j) P(\xi = x_j),$$

feltéve, hogy a $\sum\limits_{j=1}^{\infty} g(x_j) P(\xi = x_j)$ összeg abszolut konvergens.

Tétel. *Legyenek $\xi_1, \xi_2, \ldots, \xi_n$ független, diszkrét eloszlású valószínűségi változók, amelyek mindegyikére létezik az $E\xi_j$, $1 \le j \le n$, várható érték. Ekkor az $E\xi_1\xi_2 \cdots \xi_n$ várható érték is létezik, és*

$$E\xi_1\xi_2 \cdots \xi_n = E\xi_1 E\xi_2 \cdots E\xi_n.$$

A szórásnégyzet definiciója: *Legyen ξ olyan valószínűségi változó, amelyre $E\xi^2 < \infty$. Ekkor a ξ valószínűségi változó szórásnégyzetét a*

$$\mathrm{Var}\,\xi = E\left(\xi - E\xi\right)^2$$

képlet definiálja. Ha $E\xi^2 = \infty$ akkor a $\mathrm{Var}\,\xi$ szórásnégyzetet nem definiáljuk vagy azt mondjuk, hogy $\mathrm{Var}\,\xi = \infty$.

Lemma.

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

Lemma. *Minden a és b valós számra*

$$\mathrm{Var}\,(a\xi + b) = a^2 \mathrm{Var}\,\xi.$$

Tétel. Ha $\xi_1, \xi_2, \ldots, \xi_n$ *független valószínűségi változók, akkor*

$$\mathrm{Var}\,(\xi_1 + \xi_2 + \cdots + \xi_n) = \mathrm{Var}\,\xi_1 + \mathrm{Var}\,\xi_2 + \cdots + \mathrm{Var}\,\xi_n.$$

Megjegyzés: A fenti tétel fontos feltétele volt, hogy a tekintett összeg tagjai független valószínűségi változók. Ezt a feltételt lehet gyengíteni, de teljesen elhagyni nem lehet.

A kovarianciafüggvény definiciója. *Legyen ξ és η két valószínűségi változó, amelyek mindegyikének létezik szórásnégyzete, azaz $E\xi^2 < \infty$ és $E\eta^2 < \infty$. Ekkor a ξ és η kovariancifüggvényét* $\mathrm{Cov}\,(\xi,\eta)$*-t a*

$$\mathrm{Cov}\,(\xi,\eta) = E\left[(\xi - E\xi)(\eta - E\eta)\right]$$

kifejezés definiálja. Ha a $E\xi^2 < \infty$ és $E\eta^2 < \infty$ feltételek valamelyike nem teljesül, akkor nem definiáljuk a $\mathrm{Cov}\,(\xi,\eta)$ *kovariancifüggvényt.*

Lemma.

$$\mathrm{Cov}\,(\xi,\eta) = E\xi\eta - E\xi E\eta.$$

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

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

Bizonyítás.

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

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