További kiegészítő ismeretek.

Először felidézek egy eredményt arról, hogy hogyan lehet kiszámolni több diszkrét eloszlású valószínűségi változó függvényének a várható értékét. Ez hasznos lehet kovariancia (és korreláció) kiszámolásánál.

Tétel két valószínűségi változó függvényének a várható értékéről és annak kiszámolásáról. *Legyen adva két ξ és η diszkrét eloszlású valószínűségi változó, amelyeknek ismerjük az együttes eloszlásfüggvényét. Nevezetesen vegyenek fel a ξ és η valószínűségi változók valamely $x_1,\dots,x_m$ illetve $y_1,\dots,y_n$ értékeket, és legyen $P(\xi = x_j, \eta = y_k) = p_{j,k}$, $1 \le j \le m$, $1 \le k \le n$, $\sum\limits_{j=1}^{m}\sum\limits_{k=1}^{n} p_{j,k} = 1$. Legyen adva valamely két változós $f(x,y)$ függvény, és tekintsük az $f(\xi,\eta)$ valószínűségi változó várható értékét. Az*

$$Ef(\xi,\eta) = \sum_{j=1}^{m}\sum_{k=1}^{n} f(x_j,y_k)P(\xi = x_j, \eta = y_k)$$

azonosság érvényes. Speciálisan, ha a ξ és η valószínűségi változók által felvett x_1, ... , x_m és $y_1,\dots,y_n$ értékek valós számok, akkor

$$E\xi\eta = \sum_{j=1}^{m}\sum_{k=1}^{n} x_j y_k P(\xi = x_j, \eta = y_k).$$

(Hasonló képlet érvényes több valószínűségi változó függvényének a várható értékéről is, de ezt nem fogalmazom meg, mert nem lesz rá szükségünk.)

Fontos megtanulni, hogyan lehet kiszámítani általános valószínűségi változók várható értékét és szórásnégyzetét. Ilyen jellegű feladatokat fogunk tárgyalni. De ennek érdekében először meg kell értenünk néhány fontos fogalmat illetve néhány fontos eredmény jelentését. Ezért felelevenítek bizonyos az előadáson már tárgyalt fontos ismeretet. A várható érték tárgyalása előtt meg kell ismernünk az eloszlásfüggvény fogalmát. Meg kell értenünk azt, hogy hogyan tudjuk kiszámolni a várható értéket az eloszlásfüggvény segítségével. Később rátérünk a sűrűségfüggvény fogalmának ismertetésére is, és arra, hogy hogyan tudjuk kiszámolni a várható értéket a sűrűségfüggvény ismeretében.

Valószínűségi változó eloszlásfüggvényének a definiciója. *Legyen adva egy $\xi(\omega)$ (valós értékű) valószínűségi változó egy $(\Omega,\mathcal{A},P)$ valószínűségi mezőn. Ennek $F(x)$ eloszlásfüggvényén az $F(x) = P(\{\omega\colon \xi(\omega) < x\})$, $-\infty < x < \infty$, függvényt értjük.*

Tétel. *Legyen $\xi(\omega)$ valószínűségi változó egy $(\Omega,\mathcal{A},P)$ valószínűségi mezőn, és legyen $F(x) = P(\{\omega\colon \xi(\omega) < x\})$, $-\infty < x < \infty$, e valószínűségi változó eloszlásfüggvénye. Az $F(x)$ eloszlásfüggvény teljesíti a következő tulajdonságokat.*

a) $F(x)$ monoton növekvő függvény.

b) $F(x)$ balról folytonos függvény, azaz $\lim\limits_{h\to 0,\, h>0} F(x-h) = F(x)$ minden $-\infty < x < \infty$ számra.

c) $\lim_{x\to\infty} F(x) = 1.$

d) $\lim_{x\to-\infty} F(x) = 0.$

Megfordítva, ha egy $F(x)$ függvény teljesíti az előbb megfogalmazott a)—d) tulajdonságokat, akkor az eloszlásfüggvény. Részletesebben megfogalmazva ez azt jelenti, hogy ha az $F(x)$ függvény teljesíti az a)—d) feltételeket, akkor létezik egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező, és azon olyan $\xi(\omega)$ valószínűségi változó, amelynek az eloszlásfüggvénye ez az $F(x)$ függvény.

Megjegyzés. Egy ξ valószínűségi változó $F(x) = P(\xi < x)$ eloszlásfüggvénye meghatározza sok egyéb esemény valószínűségét is. Így például $P(a \le \xi < b) = P(\xi < b) - P(\xi < a) = F(b) - F(a)$, $P(a \le \xi \le b) = \lim_{\varepsilon\to 0} P(a \le \xi \le b + \varepsilon) = \lim_{\varepsilon\to 0} F(b + \varepsilon) - F(a) = F(b+0) - F(a)$. Be lehet látni, hogy a számegyenes minden szép (azaz Borel mérhető) B részhalmazára a ξ valószínűségi változó $F(x)$ eloszlásfüggvényemeghatározza a $P(\xi \in B)$ valószínűséget. Ezt a tényt úgy is interpretálhatjuk, hogy egy valószínűségi változó eloszlásfüggvénye teljes információt ad a valószínűségi változó viselkedéséről.

Valószínűségi változó várható értékének *formális* **definiciója.** *Legyen $\xi(\omega)$ valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. E valószínűségi változó várható értékét úgy definiáljuk, mint a $\xi(\omega)$ függvénynek az*

$$E\xi(\omega) = \int \xi(\omega) dP(\omega)$$

Lebesgue integrálját a P valószínűségi mérték szerint, feltéve, hogy $\int |\xi(\omega)| dP(\omega) < \infty$. Ha ez a feltétel nem teljesül, akkor nem definiáljuk a ξ valószínűségi változó várható értékét.

A Lebesgue integrál fogalmának pontos ismerete nem szükséges a továbbiak megértéséhez és a tárgyalt feladatok megoldásához. Érdemes megjegyezni, hogy ez egy a valószínűségi mezőn értelmezett függvénynek azaz valószínűségi változónak a valószínűségi mezőn definiált valószínűségi mérték szerinti integrálja. Ez az integrálfogalom hasonló a tanult Lebesgue integrál fogalmához. Először egyszerű (véges sok értéket felvevő) függvények integrálát definiáljuk természetes módon, majd természetes határátmenet segítségével ezt az integrált értelmezzük általános függvények esetén is. A defició fogalmilag nem nehéz, de a részletes elmélet kidolgozása mély mértékelméleti eredmények ismeretét igényli. Erre nem lesz szükségünk, viszont nagyon fontos tudni, hogy noha a várható érték fogalmát a valószínűségi mező és a rajta definiált valószínűségi változó és valószínűségi mérték segítségével definiáltuk, ahhoz, hogy kiszámoljuk elegendő a valószínűségi változó eloszlásának az ismerete. Ezt az alábbi jelentősége miatt *fontos tételnek* nevezett eredmény segítségével tehetjük meg.

Fontos Tétel. *Legyen $\xi(\omega)$ valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, amelynek $F(x)$ az $F(x) = P(\xi < x)$, $-\infty < x < \infty$, eloszlásfüggvénye. Ekkor*

$$E\xi(\omega) = \int_{-\infty}^{\infty} x F(\,dx),$$

ahol a fenti integrál Lebesgue–Stieltjes integrált jelöl az F mérték szerint. Ez a formula akkor érvényes, ha $\int_{-\infty}^{\infty} |x| F(\,dx) < \infty$. Ellenkező esetben az $E\xi$ várható értéket nem definiáltuk.

Sőt, igaz a Fontos Tétel következő általánosítása.

A Fontos Tétel általánosítása. *Legyen $\xi(\omega)$ valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, amelynek $F(x)$ az $F(x) = P(\xi < x)$, $-\infty < x < \infty$, eloszlásfüggvénye. Legyen $g(x)$ tetszőleges (mérhető) függvény a számegyenesen, és definiáljuk az $\eta(\omega) = g(\xi(\omega))$ valószínűségi változót. Ekkor*

$$E\eta(\omega) = Eg(\xi(\omega)) = \int_{-\infty}^{\infty} g(x) F(\,dx),$$

ahol a fenti integrál Lebesgue–Stieltjes integrált jelöl az F mérték szerint. Ez a formula akkor érvényes, ha $\int_{-\infty}^{\infty} |g(x)| F(\,dx) < \infty$. Ellenkező esetben az $E\eta$ várható értéket nem definiáltuk.

Tétel. *Ha két ξ_1, ξ_2 valószínűségi változónak (amelyek ugyanazon a valószínűségi mezőn vannak definiálva) létezik várható értéke, c_1 és c_2 két valós szám, akkor a $c_1\xi_1 + c_2\xi_2$ kifejezésnek is létezik várható értéke, és*

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

Szórásnégyzet definiciója. *Legyen ξ olyan valószínűségi változó, amelyre $E\xi^2 < \infty$. Ekkor e valószínűségi változó szórásnégyzete*

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

(Ha $E\xi^2 = \infty$, akkor vagy nem definiáljuk a ξ valószínűségi változó szórásnégyzetét, vagy azt mondjuk, hogy $\operatorname{Var}\xi(\omega) = \infty$.)

Lemma.

$$\operatorname{Var}\xi = E\xi^2 - \left(E\xi\right)^2.$$

Sűrűségfüggvény definiciója. *Egy $F(x)$ eloszlásfüggvénynek az $f(u)$, $-\infty < u < \infty$, függvény sűrűségfüggvénye, ha minden $-\infty < x < \infty$ számra teljesül az*

$$F(x) = \int_{-\infty}^{x} f(u)\,du$$

azonosság.

Megjegyzés: Többször fogunk beszélni egy valószínűségi változó sűrűségfüggvényéről is. Ez azt jelenti, hogy e valószínűségi változó eloszlásfüggvényének a sűrűségfüggvényét tekintjük. Sok érdekes és fontos eloszlásfüggvénynek létezik sűrűségfüggvénye, de nem mindegyiknek. Például egy diszkrét eloszlású valószínűségi változónak nincs sűrűségfüggvénye.

Tétel. *Ha egy $F(x)$ eloszlásfüggvénynek létezik $f(u)$ sűrűségfüggvénye, akkor ez teljesíti minden $g(\cdot)$ függvényre a következő azonosságot:*

$$\int_{-\infty}^{\infty} g(u)F(\,du) = \int_{-\infty}^{\infty} g(u)f(u)\,du.$$

Newton–Leibniz formula. *Legyen $F(x)$ folytonos függvény egy $[a,b]$ véges intervallumon, amely véges sok pont kivételével folytonosan deriválható. Legyen $f(u) = \left.\frac{dF(x)}{dx}\right|_{x=u}$ minden pontban, ahol az $F(\cdot)$ differenciálható. Ekkor $F(x) - F(a) = \int_a^x f(u)\,du$ minden $a \le x \le b$ számra. Továbbá, ha a fenti feltétel teljesül minden véges $[a,b]$ intervallumban, és létezik a $F(-\infty) = \lim\limits_{x\to-\infty} F(x)$ határérték, akkor $F(x) - F(-\infty) = \int_{-\infty}^x f(u)\,du$ minden $-\infty < x < \infty$ számra.*

Megfordítva, ha $f(u)$, (pontosabban annak abszolut értéke) integrálható függvény egy $[a,b]$ intervallumban, és $F(x) = \int_a^x f(u)\,du$, $a \le x \le b$, akkor $f(u) = \left.\frac{dF(x)}{dx}\right|_{x=u}$ (majdnem) minden $a \le u \le b$ pontban. Ha az $f(u)$ függvény integrálható az egész számegyenesen, akkor a fenti állítás igaz $a = -\infty$ választással is.

Megjegyzés: A Newton–Leibniz formula jelentősége számunkra az, hogy eszerint az eredmény szerint a sűrűségfüggvény (feltéve, hogy az létezik) egyenlő az eloszlásfüggvény deriváltjával. Megfordítva, a sűrűségfüggvény ismeretében meg tudjuk határozni az eloszlásfüggvényt. Annak értéke az x pontban megegyezik a sűrűségfüggvény integráljával a $[-\infty, x]$ intervallumban.

Számunkra a Newton–Leibniz formulának az alábbi következménye lesz fontos.

Newton–Leibniz formula következménye. *Legyen ξ olyan valószínűségi változó, amelynek nemcsak $F(x)$ eloszlás hanem $f(x)$ sűrűségfüggvénye is létezik. Ha ismerjük az $F(x)$ eloszlás $f(x)$ és sűrűségfüggvény egyikét, akkor a másik a következő képlet segítségével számolható ki.*

$$f(x) = \frac{dF(x)}{dx}, \qquad F(x) = \int_{-\infty}^{x} f(u)\,du, \qquad\qquad -\infty < x < \infty.$$

Ha a ξ valószínűségi változó eloszlásfüggvénye minden pontban folytonos, és esetleg véges sok pont kivételével deriválható is, akkor a ξ valószínűségi változónak létezik sűrűségfüggvénye is.

Tétel. *Egy $f(\cdot)$ (mérhető) függvény akkor és csak akkor sűrűségfüggvénye egy alkalmas eloszlásfüggvénynek, ha $f(u) \ge 0$ majdnem minden $-\infty < u < \infty$ számra, és $\int_{-\infty}^{\infty} f(u)\,du = 1$.*

Tétel. *Jelölje $F(\cdot)$ illetve $f(\cdot)$ egy ξ valószínűségi változó eloszlás illetve sűrűségfüggvényét. Legyen $g(\cdot)$ mérhető függvény a számegyenesen. Ekkor*

$$E\xi = \int_{-\infty}^{\infty} uF(\,du) = \int_{-\infty}^{\infty} uf(u)\,du$$

és

$$Eg(\xi) = \int_{-\infty}^{\infty} g(u)F(\,du) = \int_{-\infty}^{\infty} g(u)f(u)\,du.$$

Néhány fontos folytonos eloszlás.

a.) *Normális eloszlásfüggvény.*

Bevezetjük a standard normális eloszlásfüggvény definicióját. Későbbi eredményekből fog kiderülni, hogy ez az eloszlás miért játszik fontos szerepet a valószínűségszámításban.

A standard normális eloszlás definiciója. *A $\Phi(x)$ standard normális eloszlás az az eloszlás, amelynek sűrűségfüggvénye a $\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}$, $-\infty < u < \infty$, függvény.*

E definició helyességének igazolásához meg kell mutatni, hogy a fent definiált $\varphi(\cdot)$ függvény valóban sűrűségfüggvény. Ennek érdekében be kell bizonyítani a következő eredményt.

Tétel.

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

Normális eloszlásfüggvény definiciója: *Legyen ξ standard normális eloszlású valószínűségi változó. Ekkor az $\eta = \sigma\xi + m$ valószínűségi változó, (illetve egy vele azonos eloszlású valószínűségi változó) normális eloszlású valószínűségi változó m és σ^2 paraméterrel. (Később meg fogjuk tárgyalni, hogy m az η valószínűségi változó várható értéke és σ^2 a szórásnégyzete.)*

Egyenletes eloszlásfüggvény definiciója. *Egy ξ valószínűségi változó egyenletes eloszlású egy $[a, b]$ intervallumban, $-\infty < a < b < \infty$, ha sűrűségfüggvénye $f(u) = \frac{1}{b-a}$, ha $a \le u \le b$, és $f(u) = 0$ egyébként.*

Exponenciális eloszlásfüggvény definiciója. *Egy ξ valószínűségi változó exponenciális eloszlású λ paraméterrel, $\lambda > 0$, ha eloszlásfüggvénye, $F(x) = P(\xi < x) = 1 - e^{-\lambda x}$, ha $x \ge 0$, és $F(x) = P(\xi < x) = 0$, ha $x < 0$. Ezzel ekvivalens jellemzés: Egy valószínűségi változó exponenciális eloszlású $\lambda > 0$ paraméterrel, ha létezik $f(u)$ sűrűségfüggvénye, és az $f(u) = \lambda e^{-\lambda u}$, ha $u \ge 0$, $f(u) = 0$, ha $u \le 0$ alakú.*

Többdimenziós eloszlásfüggvény definiciója. *Legyen adva k valós értékű $\xi_1, \dots, \xi_k$ valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Ezek (együttes) eloszlásfüggvénye az*

$$F(x_1, \dots, x_k) = P(\xi_1 < x_1, \dots, \xi_k < x_k),$$

k változós függvény, ahol $-\infty < x_j < \infty$ minden $1 \le j \le k$ indexre.

Hasonlóan az egyváltozós eloszlásfüggvényekhez hasonlóan lehet jellemezni a többdimenziós eloszlásfüggvényeket is. Erről szól a következő tétel.

Tétel. *Egy $F(u_1, \dots, u_k)$ függvény akkor és csak akkor eloszlásfüggvénye alkalmas $\xi_1, \dots, \xi_k$ valószínűségi változóknak valamely $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, ha ez az F függvény teljesíti a következő (i)—(iv) tulajdonságokat.*

(i) $F(u_1, \dots, u_k)$ minden változójának balról folytonos függvénye.

(ii) $\lim\limits_{\substack{u_j \to \infty \\ \text{minden } j=1,\dots,k \text{ számra}}} F(u_1, \dots, u_k) = 1.$

(iii) $\lim\limits_{\substack{u_j \to -\infty \\ \text{valamely } 1 \le j \le k \text{ számrra}}} F(u_1, \dots, u_k) = 0.$ *(Ez úgy értendő, hogy az összes u_s, $1 \le s \le k$, $s \ne j$ koordinátát rögzítjük, és $u_j \to -\infty$.)*

Végül definiáljuk egy az R^k téren definiált F függvényre és egy $\mathbf{K} = [a_1, b_1) \times \cdots \times [a_k, b_k)$ téglatestre a

$$\mu(\mathbf{K}) = \mu_F(\mathbf{K}) = \sum_{\substack{u_j = \, a_j \text{ vagy } b_j \\ j=1,\dots,k}} (-1)^{\chi(u_1,\dots,u_k)} F(u_1, \dots, u_k)$$

mennyiséget, ahol $\chi(u_1, \dots, u_k)$ jelöli az a_j-k számát az $u_1, \dots, u_k$ sorozatban. Ekkor

(iv) $\mu_F(\mathbf{K}) \ge 0$ minden $\mathbf{K}$ téglatestre.

Hasonlóan az egydimenziós eloszlás sűrűségfüggvényéhez definiálhatjuk a többdimenziós eloszlás sűrűségfüggvényét is. Az egyváltozós esethez hasonlóan ennek is megadhatjuk a jellemzését. Erről szól az alábbi definició és tétel.

Több-dimenziós eloszlásfüggvény sűrűségfüggvényének definiciója. *Azt mondjuk, hogy egy $F(x_1, \dots, x_k)$ többváltozós eloszlásfüggvénynek létezik $f(u_1, \dots, u_k)$ sűrűségfüggvénye, ha teljesül az*

$$F(x_1, \dots, x_k) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_k} f(u_1, \dots, u_k)\, du_1 \dots du_k$$

azonosság minden $-\infty < x_j < \infty$, $1 \le j \le k$ számra.

Tétel. *Egy k-változós $f(u_1, \dots, u_k)$ függvény akkor és csak akkor sűrűségfüggvénye egy alkalmas k-dimenziós eloszlásfüggvénynek, ha $f(u_1, \dots, u_k) \ge 0$ majdnem minden $(u_1, \dots, u_k)$ pontban, és*

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(u_1, \dots, u_k)\, du_1 \dots du_k = 1.$$

A többdimenziós eloszlásfüggvény fogalmának a segítségével definiálhatjuk valószínűségi változók függetlenségét az általános esetben is. Kimondok egy a független valószínűségi változók fontos tulajdonságát kimondó tételt. Hangsúlyozom, hogy korábban csak diszkrét eloszlású valószínűségi változók függetlenségéről beszéltünk. Hasznos megérteni a két fogalom kapcsolatát is. Ebben is segít az alább ismertetett tétel.

Valószínűségi változók függetlenségének a definiciója. *Legyenek $\xi_1, \dots, \xi_n$ valószínűségi változók egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Azt mondjuk, hogy ezek a valószínűségi változók függetlenek, ha minden $x_1, \dots, x_n$ valós számra*

$$P(\xi_1 < x_1, \dots, \xi_n < x_n) = P(\xi_1 < x_1) \cdots P(\xi_n < x_n).$$

Tétel. *Legyenek $\xi_1, \dots, \xi_k$ független valószínűségi változók, $B_1, \dots, B_k$ a számegyenes Borel mérhető részhalmazai. Ekkor*

$$P(\xi_1 \in B_1, \dots, \xi_k \in B_k) = P(\xi_1 \in B_1) \cdots P(\xi_k \in B_k).$$

Ugyancsak fontos a következő eredmény, amely független valószínűségi változók szorzatának a várható értékéről szól. Korábban ennek az eredménynek csak a független, diszkrét eloszlású valószínűségi változók szorzatáról szóló speciális esetét tárgyaltuk.

Tétel. *Legyenek $\xi_1, \dots, \xi_k$ független valószínűségi változók, amelyek mindegyikének létezik várható értéke, azaz $E|\xi_j| < \infty$. Ekkor a $\xi_1 \cdots \xi_k$ szorzatnak is létezik várható értéke, és*

$$E\xi_1 \cdots \xi_k = E\xi_1 \cdots E\xi_k.$$

További hasznos eredmények:

Tétel. *Ha $\xi_1, \dots, \xi_k$ független valószínűségi változók, $h_1(x), \dots, h_k(x)$ tetszőleges (mérhető) függvények, akkor $\eta_1 = h_1(\xi_1)$, $\dots$, $\eta_k = h_k(\xi_k)$ szintén független valószínűségi változók.*

Tétel. *Ha $\xi_1, \dots, \xi_k$ független valószínűségi változók, $c_1, \dots, c_k$ tetszőleges valós számok, akkor*

$$\mathrm{Var}\,(c_1\xi_1 + \dots + c_k\xi_k) = c_1^2 \mathrm{Var}\,\xi_1 + \dots + c_k^2 \mathrm{Var}\,\xi_k.$$

Tétel. *Legyenek ξ_j, $1 \le j \le k$, valószínűségi változók ugyanazon a valószínűségi mezőn, amelyekre teljesül $E\xi_j^2 < \infty$ feltétel. Ekkor*

$$\mathrm{Var}\left(\sum_{j=1}^k \xi_j\right) = \sum_{j=1}^k \mathrm{Var}\,\xi_j + \sum_{\substack{1 \le j,l \le k \\ j \ne l}} \mathrm{Cov}\,(\xi_j, \xi_l) = \sum_{j=1}^k \mathrm{Var}\,\xi_j + 2\sum_{j=1}^{k-1} \sum_{l=j+1}^{k} \mathrm{Cov}\,(\xi_j, \xi_l).$$

Speciálisan, ha a ξ_j valószínűségi változók függetlenek, akkor

$$\mathrm{Var}\left(\sum_{j=1}^k \xi_j\right) = \sum_{j=1}^k \mathrm{Var}\,\xi_j.$$

Az alábbi eredmény szintén fontos a valószínűségszámításban, de időhiány miatt valószínűleg nem tudjuk tárgyalni, és lehet, hogy ugyancsak időhiány miatt az előadáson sem kerül sor ennek tárgyalására. A biztonság kedvéért mégis közlöm.

Tétel független valószínűségi változók összegének a sűrűségfüggvényéről. *Legyen ξ és η két független valószínűségi változó $f(\cdot)$ és $g(\cdot)$ sűrűségfüggvénnyel. Ekkor a $\xi+\eta$ összegnek is létezik sűrűségfüggvénye, és az az*

$$f * g(x) = \int_{-\infty}^{\infty} f(u)g(x-u)\,du, \quad -\infty < x < \infty,$$

függvény.