Valószínűségi döntéstámogató rendszerek

Antos András – Antal Péter – Hullám Gábor – Millinghoffer András – Hajós Gergely

Kulcsszavak:
döntés, becslés, költségfüggvény, kockázat, a priori és a posteriori valószínűség, Bayes-döntés és -becslés, Bayes-statisztika, valószínűségi gráf alapú modellek, Bayes-háló, rejtett Markov modell, emberi becslési heurisztikák, valószínűségi következtetés, bayesi döntéselmélet, optimális döntés, információ értéke, többkarú rabló probléma, QUALY, költség-haszon elemzés, döntési hálók

Összefoglalás:
A jegyzetben ismertetjük a döntés- és becsléselmélet alapfogalmait és a leggyakoribb költségfüggvényeket. Megvizsgáljuk a Bayes-döntést, maximum a posteriori és maximum likelihood döntést és a Bayes-döntés közelítését több példán keresztül. Kitérünk a Bayes-becslésre, maximum likelihood becslésre és regressziós becslésre részletesen megvizsgálva a lineáris regresszió esetét. Ezt követően a valószínűségi eloszlások strukturális jellemzőit vizsgáljuk meg. A valószínűségi gráfos modellosztályon belül elsőként az egyszerű Naív Bayes-háló, Markov-lánc és rejtett Markov modell modelltípusokat foglaljuk össze, majd a Bayes-hálókat és a Markov-hálókat tekintjük át.

Budapesti Műszaki és Gazdaságtudományi Egyetem
Semmelweis Egyetem

Typotex Kiadó
2014

Szerkesztette: Antal Péter
Szakmai lektor: Kovács András

ISBN 978 963 279 184 5

Készült a Typotex Kiadó (`http://www.typotex.hu`) gondozásában
Felelős vezető: Votisky Zsuzsa

Készült a TÁMOP-4.1.2.A/1-11/1-2011-0079 számú, „Konzorcium a biotechnológia és bioinformatika aktív tanulásáért” című projekt keretében

Tartalomjegyzék

1. fejezet

Valószínűségi becslés- és döntéselmélet

1.1. Bevezetés

Ez a fejezet alapvetően a [4] jegyzet 5. fejezetén alapszik. Továbbá az 1.2, 1.3, 1.4 és 1.5 fejezetekhez hasznos lehet a [2] könyv 1. és 2. fejezete, az 1.2, 1.6, 1.7 és 1.8 fejezetekhez pedig a [3] könyv 1. fejezete.

1.2. Definíciók

Gyakori probléma, hogy egy X megfigyelhető mennyiségből kell egy másik, közvetlenül (még) nem megfigyelhető Y mennyiségre következtetnünk, annak értéket megbecsülnünk. Általános esetben matematikailag mindkét mennyiséget valószínűségi változókkal modellezhetjük. Jelölje X, illetve Y értékkészletét (azaz lehetséges értékeinek halmazát) $\mathcal{X}$, illetve $\mathcal{Y}$. Ekkor formálisan $X : \Omega \to \mathcal{X}$, $Y : \Omega \to \mathcal{Y}$ függvények (ahol Ω a valószínűségi mező alaphalmaza).

1. példa. $\mathcal{X}$ és $\mathcal{Y}$ tipikusan lehet például $\mathbf{R}$ (a valós számok halmaza), $\mathbf{R}^d$, $\{0,1\}^d$, $[0,1]^d$, vagy ezek tetszőleges megszámlálható, esetleg véges részhalmaza. Legegyszerűbb esetben $\mathcal{Y} = \{0,1\}$.

X értékéből Y-t a $g : \mathcal{X} \to \bar{\mathcal{Y}}$ *következtetésfüggvénnyel* próbáljuk meghatározni, ahol $\bar{\mathcal{Y}}$ az elképzelhető következtetésfüggvények értékkészletének uniója. $g(X)$-t *következtetésnek* nevezzük. Tipikusan $\bar{\mathcal{Y}} \supseteq \mathcal{Y}$. Egy következtetés jóságát egy nemnegatív $C : \mathcal{Y} \times \bar{\mathcal{Y}} \to [0, \infty)$ *költségfüggvény* (vagy *jósági*, *hasonlósági*, *megbízhatósági kritérium*) méri, azaz $C(y, y')$ a költsége annak, ha a valódi y érték helyett y'-re következtetünk. Például ha a megfigyelés x volt, és $Y = y$, akkor a g által adott következtetés költsége $C(y, g(x))$. Minél kisebb ez a költség, a következtetés annál jobbnak tekinthető.
Ha minden $y \in \mathcal{Y}$-ra $C(y, y')$ konstans minden $y' \in \bar{\mathcal{Y}} \setminus \{y\}$-ra, azaz pontatlan következtetés esetén mindig azonos, akkor *döntési problémáról* beszélünk. Ilyenkor el akarjuk találni y-t. Ha nem találjuk el, mindegy, hogy „mennyire" nem. Ha C valamiféle intuitív távolság, akkor *becslési problémáról* beszélünk. Ilyenkor az is számít, hogy mennyit tévedünk. Döntési problémáknál általában $\bar{\mathcal{Y}} = \mathcal{Y}$ véges vagy megszámlálható (diszkrét), míg becslési problémáknál általában $\bar{\mathcal{Y}}$ vagy akár $\mathcal{Y}$ is folytonos (megszámlálhatatlan végtelen).

Mivel $C(Y, g(X))$ függ (X, Y)-tól, maga is valószínűségi változó. Így g jóságát a várható költsége, az

$$R(g) \stackrel{\text{def}}{=} \mathbf{E}\left[C(Y, g(X))\right]$$

globális kockázat (risk) méri. Célunk annak a következtetésfüggvénynek a megtalálása, amelyre $R(g)$ a legkisebb. Ezt *Bayes-feladatnak*, míg az ilyen g-t *optimálisnak* nevezzük.
Legyen

$$r(g, x) \stackrel{\text{def}}{=} \mathbf{E}\left[C(Y, g(X))|X = x\right]$$

a *lokális kockázat függvény*, azaz g költségének feltételes várhatóértéke $X = x$ esetén.

1. gyakorlat. Lássuk be, hogy $g(X)$ helyett írhatunk $g(x)$-et is.

Nyilván a teljes várhatóérték tétele és a feltételes várhatóérték definíciója szerint

$$\{eq : explok_g lob\mathbf{E}\left[r(g, X)\right] = R(g) \tag{1.1}$$

és

$$r(g, x) = \mathbf{E}\left[C(Y, g(x))|X = x\right] = \int_{\mathcal{Y}} C(y, g(x))\, dF_{Y|x}(y), \tag{1.2}$$

ahol $F_{Y|x}$ az Y feltételes eloszlásfüggvénye, ha $X = x$, amelyet *a posteriori eloszlásfüggvénynek* is nevezünk. Ez tehát általános esetben ún. (Lebesgue-)Stieltjes-integrállal írható fel, de a továbbiakban általunk vizsgált diszkrét, illetve abszolút folytonos eloszlásokra egyszerű összegzésre, illetve Riemann-integrálra vezet. Az a posteriori eloszlás elnevezés arra utal, hogy ez Y eloszlása X értékének ismerete *után*. Ettől megkülönböztetendő Y feltétel nélküli, eredeti (azaz X értékének ismerete *előtti*) eloszlását az Y *a priori eloszlásának* is nevezzük.

1.2.1. Gyakori költségfüggvények és tulajdonságaik

Döntési problémáknál a leggyakoribb költségfüggvény választás a következő:

2. példa. *0-1* költség: Legyen $C_0(y, y') \stackrel{\text{def}}{=} \mathbf{I}_{\{y \neq y'\}}$ (ahol $\mathbf{I}_A$ az A esemény indikátorfüggvénye, azaz $\mathbf{I}_A = 1$, ha A bekövetkezik és $\mathbf{I}_A = 0$, ha nem). Ekkor

$$R(g) = \mathbf{E}\left[C_0(Y, g(X))\right] = \mathbf{E}\left[\mathbf{I}_{\{Y \neq g(X)\}}\right] = \mathbf{P}\left[Y \neq g(X)\right],$$

vagyis a globális kockázat éppen g hibázásának a valószínűsége.

Becslési problémáknál $\bar{\mathcal{Y}} = \mathcal{Y} = \mathbf{R}^d$-re gyakori választás az L_p távolság (más jelöléssel a $\|\cdot\|_p$ norma) p-edik hatványa, azaz $C_p(y, y') \stackrel{\text{def}}{=} \sum_{s=1}^{d} |y_s - y'_s|^p = \|y - y'\|_p^p$. Ekkor

$$R(g) = \mathbf{E}\left[C_p(Y, g(X))\right] = \mathbf{E}\left[\|Y - g(X)\|_p^p\right].$$

Leggyakoribb a $p = 1$ és 2 eset (ezért $\|\cdot\|_1$ helyett $\|\cdot\|$-t is használunk):

3. példa. $C_1(y, y') = \|y - y'\|$ az *abszolút* költség. Ekkor például $\bar{\mathcal{Y}} = \mathbf{R}$-re $R(g) = \mathbf{E}\left[|Y - g(X)|\right]$ az *abszolút középhiba*.

4. példa. $C_2(y,y') = \|y-y'\|_2^2$ a *négyzetes* költség. Ekkor például $\bar{\mathcal{Y}} = \mathbf{R}$-re $R(g) = \mathbf{E}\left[(Y-g(X))^2\right]$ a *négyzetes középhiba.*

Érdemes először megvizsgálnunk, hogy a fenti példák – és az egyszerűség kedvéért 2.-ben diszkrét $\mathcal{Y}$, 3., 4.-ben $\bar{\mathcal{Y}} = \mathbf{R}$ – esetén milyen $g \in \bar{\mathcal{Y}}$ érték minimalizálja a $\mathbf{E}\left[C(Y,g)\right]$ várható költséget, azaz rendre

$$\begin{aligned}
&\mathbf{E}\left[C_0(Y,g)\right] = \mathbf{P}\left[Y \neq g\right], \\
&\mathbf{E}\left[C_1(Y,g)\right] = \mathbf{E}\left[|Y-g|\right], \\
&\mathbf{E}\left[C_2(Y,g)\right] = \mathbf{E}\left[(Y-g)^2\right] \text{-t.}
\end{aligned}$$

$\mathbf{P}\left[Y \neq g\right]$ minimalizálása nyilván $\mathbf{P}\left[Y = g\right]$ maximalizálását jelenti, vagyis g-t az

$$\arg\max_{y \in \mathcal{Y}} \mathbf{P}\left[Y = y\right]$$

halmaz egyik elemének, azaz Y (nem mindig egyértelmű) *móduszának* kell választani.

2. gyakorlat. Lássuk be, hogy $\mathbf{E}\left[|Y-g|\right]$-t Y (nem mindig egyértelmű) *mediánja* minimalizálja, azaz azon g értékek, amelyekre $\mathbf{P}\left[Y \leq g\right] \geq 1/2$ és $\mathbf{P}\left[Y \geq g\right] \geq 1/2$. (Például abszolút folytonos Y-ra ez $\mathbf{P}\left[Y \leq g\right] = \mathbf{P}\left[Y \geq g\right] = 1/2$-ként írható.)

A négyzetes középhiba minimalizálásában segít a következő

1.1. tétel (Steiner-tétel)**.** Bármely véges szórású Y valós valószínűségi változóra és $g \in \mathbf{R}$-re

$$\mathbf{E}\left[(Y-g)^2\right] = \mathbf{E}\left[(Y-\mathbf{E}\left[Y\right])^2\right] + (\mathbf{E}\left[Y\right]-g)^2.$$

Bizonyítás.

$$\begin{aligned}
\mathbf{E}\left[(Y-g)^2\right] &= \mathbf{E}\left[((Y-\mathbf{E}\left[Y\right]) + (\mathbf{E}\left[Y\right]-g))^2\right] = \\
&= \mathbf{E}\left[(Y-\mathbf{E}\left[Y\right])^2\right] + 2\mathbf{E}\left[(Y-\mathbf{E}\left[Y\right])(\mathbf{E}\left[Y\right]-g)\right] + \mathbf{E}\left[(\mathbf{E}\left[Y\right]-g)^2\right] = \\
&= \mathbf{E}\left[(Y-\mathbf{E}\left[Y\right])^2\right] + (\mathbf{E}\left[Y\right]-g)^2,
\end{aligned}$$

mivel a középső tagban $(\mathbf{E}\left[Y\right]-g)\mathbf{E}\left[Y-\mathbf{E}\left[Y\right]\right] = 0$. □

1. *megjegyzés.* A tétel ($\mathbf{R}^2$-re vonatkozó verziójának) fizikai analógiája az, hogy egy test tehetetlenségi nyomatéka egy tetszőleges tengely körül a tehetetlenségi nyomatéka a tömegközépponton átmenő, párhuzamos tengely körül, plusz a test tömege szorozva a két tengely merőleges távolságának négyzetével (párhuzamos tengely tétel). Innen származik a Steiner-tétel elnevezés.

A 1.1. tétel következménye, hogy

$$\min_{g \in \mathbf{R}} \mathbf{E}\left[(Y-g)^2\right] = \mathbf{E}\left[(Y-\mathbf{E}\left[Y\right])^2\right],$$

azaz $\mathbf{E}\left[(Y-g)^2\right]$-t a $g = \mathbf{E}\left[Y\right]$ választás minimalizálja.
Összefoglalva, a minimalizáló értékek rendre a következők:

$$\begin{aligned}
&C_0 \text{ :módusz } \left(\arg\max_{y \in \mathcal{Y}} \mathbf{P}\left[Y = y\right]\right), \\
&C_1 \text{ :medián,} \\
&C_2 \text{ :várhatóérték (mean) } \left(\mathbf{E}\left[Y\right]\right).
\end{aligned}$$

Látni fogjuk, hogy amikor $g = g(X)$-t az X-től függően választjuk, akkor ezen megfigyelések feltételes verziója lép érvénybe, azaz a minimalizáló értékek helyébe az Y feltéve X feltételes eloszlás módusa, mediánja, illetve várhatóértéke lép. A bizonyítások a fentiekkel analóg módon történnek.
Az L_p távolságon alapuló C_p költségfüggvények közötti további fontos összefüggés a következő:

1.2. tétel. $\bar{\mathcal{Y}} = \mathcal{Y} = \mathbf{R}^d$ esetén mind $(C_p(y,y')/d)^{1/p} = \|y-y'\|_p d^{-1/p}$, mind $\mathbf{E}\left[C_p(Y,g(X))/d\right]^{1/p}$ monoton növekvő p-ben. Speciálisan $\bar{\mathcal{Y}} = \mathbf{R}$-re, $C_p^{1/p}(y,y')$ és $\mathbf{E}\left[C_p(Y,g(X))\right]^{1/p}$ monoton növekvő p-ben.

Bizonyítás. Bármely $0 < p < q$-ra

$$\left(\frac{C_p(y,y')}{d}\right)^{1/p} = \left(\sum_{s=1}^{d}\frac{1}{d}|y_s-y'_s|^p\right)^{1/p} = \left(\left(\sum_{s=1}^{d}\frac{1}{d}|y_s-y'_s|^p\right)^{q/p}\right)^{1/q} \leq$$

$$\leq \left(\sum_{s=1}^{d}\frac{1}{d}(|y_s-y'_s|^p)^{q/p}\right)^{1/q} = \left(\sum_{s=1}^{d}\frac{1}{d}|y_s-y'_s|^q\right)^{1/q} = \left(\frac{C_q(y,y')}{d}\right)^{1/q},$$

ahol az egyenlőtlenség a Jensen-egyenlőtlenség alkalmazása egy egyenletes d értékű eloszlásra, hiszen $q/p > 1$, és így $x^{q/p}$ konvex függvény. A bizonyítás az (X,Y) szerinti várhatóértékkel együtt is pontosan ugyanígy történik a Jensen-egyenlőtlenség kétszeri alkalmazásával. □

Sokszor az $y \in \mathbf{R}^d$ lehetséges értékei speciálisan eloszlások súlyvektorai. Ilyenkor gyakran hasznos költségfüggvény választás az ún. *Kullback–Leibler- (KL-) távolság* (vagy *relatív entrópia*, *I-divergencia*) [1, 2.3. fejezet]:

5. példa. Legyen $C_{\mathrm{KL}}(y,y') \stackrel{\mathrm{def}}{=} D_{\mathrm{KL}}(y\|y') = \sum_{s=1}^{d} y_s \ln\frac{y_s}{y'_s}$.

Ismert, hogy a KL-távolság dominálja az L_1 távolság négyzetét [1, p. 300, Lemma 12.6.1]:

1.3. tétel (Pinsker-egyenlőtlenség)**.** $2C_{\mathrm{KL}}(y,y') \geq \|y-y'\|^2 = C_1^2(y,y')$.

1.3. Bayes-döntés

Vizsgáljuk meg először alaposabban azt az esetet, amikor $\mathcal{Y} = \bar{\mathcal{Y}}$ az $\{y_1,\ldots,y_k\}$ véges halmaz. Ekkor

1. definíció. Az $\{Y = y_i\}$ eseményt *i-edik hipotézisnek* nevezzük. Y (a priori) eloszlását a $q_i \stackrel{\mathrm{def}}{=} \mathbf{P}[Y = y_i]$ *a priori valószínűségek*, míg az a posteriori eloszlását az $\eta_i(x) \stackrel{\mathrm{def}}{=} \mathbf{P}[Y = y_i|X = x]$ *a posteriori valószínűségek* adják meg.
g-t *döntésfüggvénynek*, $g(X)$-t *döntésnek* is nevezzük. Az y_i értékek g-vel való ősképei $\mathcal{X}$ egy partícióját adják, amelynek $D_i = \{x \in \mathcal{X} : g(x) = y_i\}$ osztályait $(i = 1,2,\ldots,k)$ *döntési tartományoknak* nevezzük. D_i tehát $\mathcal{X}$ azon maximális részhalmaza, amely bármely elemének megfigyelésekor y_i-re dönt g.

3. gyakorlat. Mutassuk meg, hogy adott $\bar{\mathcal{Y}} = \{y_1, \ldots, y_k\}$-ra a $(D_1, \ldots, D_k)$ döntési tartományok teljesen meghatározzák a g döntésfüggvényt, azaz a döntésfüggvényt megadhatjuk a döntési tartományok rendezett k-asával, továbbá hogy $\mathbf{I}_{\{g(x)=y_j\}} = \mathbf{I}_{\{x\in D_j\}}$.

Most – bevezetve a $C(y_i, y_j) = C_{ij}$ rövidítést – (1.2) szerint a lokális kockázat

$$r(g,x) = \sum_{i=1}^{k} C(y_i, g(x))\eta_i(x) = \sum_{i=1}^{k}\sum_{j=1}^{k} \mathbf{I}_{\{g(x)=y_j\}} C(y_i, y_j)\eta_i(x) =$$
$$= \sum_{i=1}^{k}\sum_{j=1}^{k} \mathbf{I}_{\{x\in D_j\}} C_{ij}\eta_i(x) = \sum_{j=1}^{k} \mathbf{I}_{\{x\in D_j\}} \sum_{i=1}^{k} C_{ij}\eta_i(x) = \sum_{j=1}^{k} \mathbf{I}_{\{x\in D_j\}} d_j(x), \quad (1.3)$$

ahol $d_j(x) \stackrel{\text{def}}{=} \sum_{i=1}^{k} C_{ij}\eta_i(x)$. Ez éppen a költség feltételes várhatóértéke, ha $X = x$, *és* a következtetés a j-edik hipotézisre dönt.
Látható, hogy a fenti $r(g,x)$-et olyan g minimalizálja, amely x-et az (egyik) legkisebb $d_j(x)$-hez tartozó D_j tartományba sorolja (lásd 1.4. tétel alább). Ha több legkisebb $d_j(x)$ van, akkor mindegy, melyiket választjuk, legyen ez például legkisebb indexű. Legyenek tehát a D_j^* tartományok olyanok, hogy $\forall x \in D_j^*$ akkor és csak akkor, ha $d_j(x) < d_i(x)\ \ \forall i < j$-re és $d_j(x) \le d_i(x)\ \ \forall i > j$-re, vagyis x pontosan akkor eleme D_j^*-nak, ha az $X = x$ esetén a j-edik hipotézisre való döntés (feltételes) várható költsége a(z egyik) legkisebb, és több minimális várható költség esetén a hipotézisek indexei közül j a legkisebb.

4. gyakorlat. Mutassuk meg, hogy a fenti D_j^*-k páronként diszjunktak, és uniójuk $\mathcal{X}$, azaz $\mathcal{X}$ partícióját adják, továbbá hogy

$$x \in D_j^* \Rightarrow d_j(x) = \min_{1\le i\le k} d_i(x). \quad (1.4)$$

2. definíció. A fenti $(D_1^*, \ldots, D_k^*)$ tartományok által meghatározott g^* döntésfüggvényt (azaz amelyre $g^*(x) = y_j \Leftrightarrow x \in D_j^*$) *Bayes-döntésnek* nevezzük.

1.4. tétel. A Bayes-döntés $\forall x$-re minimalizálja a lokális kockázatot, és így optimális. A minimum értéke $r(g^*, x) = \min_{1\le i\le k} d_i(x)$.

Bizonyítás. Bármely g döntésfüggvényre és $\forall x \in \mathcal{X}$-re (1.3) szerint

$$r(g,x) = \sum_{j=1}^{k} \mathbf{I}_{\{x\in D_j\}} d_j(x) \ge \sum_{j=1}^{k} \mathbf{I}_{\{x\in D_j\}} \min_{1\le i\le k} d_i(x) = \min_{1\le i\le k} d_i(x) \sum_{j=1}^{k} \mathbf{I}_{\{x\in D_j\}} =$$
$$= \min_{1\le i\le k} d_i(x) = \min_{1\le i\le k} d_i(x) \sum_{j=1}^{k} \mathbf{I}_{\{x\in D_j^*\}} = \sum_{j=1}^{k} \mathbf{I}_{\{x\in D_j^*\}} \min_{1\le i\le k} d_i(x) =$$
$$= \sum_{j=1}^{k} \mathbf{I}_{\{x\in D_j^*\}} d_j(x) = r(g^*, x),$$

ahol az utolsó két egyenlőséghez (1.4)-t, majd ismét (1.3)-t használtuk. Tehát $r(g*, x) \le r(g,x)$, és így (**??**) alapján $R(g*) \le R(g)$, azaz g^* optimális. □

Tehát a Bayes-döntés (optimális) globális kockázata

$$R^* \stackrel{\text{def}}{=} R(g^*) = \mathbf{E}\left[\min_{1\le i\le k} d_i(X)\right]. \tag{1.5}$$

Ezt *Bayes-kockázatnak* is nevezzük.
Vizsgáljuk meg a 2. példabeli $C_{ij} = C_0(i,j) = \mathbf{I}_{\{i\neq j\}}$ költségfüggvényt. Ekkor

$$d_j(x) = \sum_{i=1}^{k} C_{ij}\eta_i(x) = \sum_{i=1}^{k} \mathbf{I}_{\{i\neq j\}}\eta_i(x) = \sum_{i=1}^{k} \eta_i(x) - \eta_j(x) = 1 - \eta_j(x).$$

Így most $x \in D_j^*$ pontosan akkor, ha $\eta_j(x) > \eta_i(x)\ \forall i < j$-re és $\eta_j(x) \ge \eta_i(x)\ \forall i > j$-re, vagyis ekkor $\eta_j(x) = \max_{1\le i\le k} \eta_i(x)$. Tehát most g^* azt a hipotézist választja, amelyik a legvalószínűbb a megfigyelés ismeretében, azaz g^*-ot a maximális a posteriori valószínűségek megkeresésével határozhatjuk meg. Ezért ekkor a Bayes-döntést *maximum a posteriori döntésnek* is nevezik.
Ebben az esetben g lokális kockázata (1.3) szerint

$$r(g,x) = \sum_{j=1}^{k} \mathbf{I}_{\{x\in D_j\}}(1-\eta_j(x)) = 1 - \sum_{j=1}^{k} \mathbf{I}_{\{x\in D_j\}}\eta_j(x),$$

ami $g = g^*$-ra a fentiek alapján így egyszerűsíthető:

$$r(g^*,x) = 1 - \sum_{j=1}^{k} \mathbf{I}_{\{x\in D_j^*\}}\eta_j(x) = 1 - \sum_{j=1}^{k} \mathbf{I}_{\{x\in D_j^*\}} \max_{1\le i\le k} \eta_i(x) = 1 - \max_{1\le i\le k} \eta_i(x).$$

Tehát a globális kockázat (**??**) alapján ekkor $R^* = R(g^*) = 1 - \mathbf{E}\left[\max_{1\le i\le k} \eta_i(X)\right]$, amit *Bayes-hibának* is nevezünk. [2, 2.1. fejezet]

6. példa. Speciálisan ha csak két hipotézis van, azaz Y bináris ($k = 2$), akkor a kockázatok egyszerűen így is írhatóak:

$$r(g^*,x) = 1 - \max(\eta_1(x),\eta_2(x)) = \min(\eta_1(x),\eta_2(x))$$

és

$$R(g^*) = \mathbf{E}\left[\min(\eta_1(X),\eta_2(X))\right] = \mathbf{E}\left[\min(\eta_1(X),1-\eta_1(X))\right].$$

Ha X diszkrét vagy abszolút folytonos változó, akkor az η_i-ket kifejezhetjük az eloszlásokból.
Legyen először X diszkrét (azaz $\mathcal{X}$ megszámlálható halmaz), és jelölje

$$p_i(x) = \mathbf{P}\left[X = x | Y = y_i\right]$$

az X feltételes súlyfüggvényét, ha $Y = y_i$. Ekkor a pozitív eséllyel előforduló x-ekre a feltételes valószínűség definíciója szerint

$$\eta_i(x) = \mathbf{P}\left[Y = y_i | X = x\right] = \frac{\mathbf{P}\left[Y = y_i, X = x\right]}{\mathbf{P}\left[X = x\right]} = \\ = \frac{\mathbf{P}\left[Y = y_i\right]\mathbf{P}\left[X = x | Y = y_i\right]}{\mathbf{P}\left[X = x\right]} = \frac{q_i p_i(x)}{\mathbf{P}\left[X = x\right]}. \tag{1.6}$$

Tehát a maximális a posteriori valószínűséget – és így a C_0 0-1 költség esetén a Bayes-döntést – az a hipotézis adja, amelyikre $q_i p_i(x)$ maximális.
Abban a speciális esetben, ha minden q_i egyenlő (szükségképpen $1/k$), azaz Y (a priori) eloszlása egyenletes, akkor $q_i p_i(x)$ és $p_i(x)$ ugyanazon i-re maximális, tehát a Bayes-döntés azt az i-t választja, amelyikre $p_i(x)$ maximális.
Amikor a q_i valószínűségek ismeretlenek, sokszor nincs jobb kiindulás, mint egyenletes a priori eloszlást feltételezni, és ezért a fentiek alapján arra a hipotézisre dönteni, amelyikre $p_i(x)$ maximális, azaz amelyiket feltételezve a megfigyelésnek a legnagyobb a valószínűsége. (Döntetlen esetén a választás most is tetszőleges lehet.)

3. definíció. Diszkrét X esetén g-t *maximum likelihood döntésnek* nevezzük, ha $g(x) = y_j$ esetén $p_j(x) = \max_i p_i(x)$.

2. *megjegyzés.* Figyeljük meg, hogy itt a maximalizálandó feltételes valószínűségben az argumentum és a feltétel éppen fel van cserélve a maximum a posteriori döntéshez képest.

Legyen most $\mathcal{X} \subseteq \mathbf{R}$, X abszolút folytonos $f(x)$ sűrűségfüggvénnyel, és jelölje

$$f_i(x) = f(x|Y = y_i)$$

az X feltételes sűrűségfüggvényét, ha $Y = y_i$. Ekkor η_i precíz definíciója technikailag bonyolultabb, de határérték számítással meggondolható, hogy $f(x) \neq 0$ esetén a diszkrét esettel és az (1.6) egyenlőséggel formailag analóg módon

$$\eta_i(x) = \frac{q_i f_i(x)}{f(x)} \tag{1.7}$$

adódik. Ekkor tehát a Bayes-döntés megkeresése – 0-1 költség esetén – $q_i f_i(x)$ maximalizálását, a maximum likelihood döntésé pedig $f_i(x)$ maximalizálását jelenti:

4. definíció. Abszolút folytonos X esetén g-t *maximum likelihood döntésnek* nevezzük, ha $g(x) = y_j$ esetén $f_j(x) = \max_i f_i(x)$.

3. *megjegyzés.* Az (1.7) egyenlőség fennállását illusztrálhatjuk egy speciális esettel: Közelítsük $\eta_i(x)$-ben az $\{X = x\}$ feltételt az $S_\epsilon(x) \stackrel{\text{def}}{=} \{x - \epsilon < X < x + \epsilon\}$ eseményekkel, ahol $\epsilon > 0$ egyre kisebb, majd $\eta_i(x)$-t a $\eta_i^\epsilon(x) \stackrel{\text{def}}{=} \mathbf{P}[Y = y_i | S_\epsilon(x)]$ függvényekkel. Ekkor, ha $f_1, \ldots, f_k$, és következésképpen $f(x)$ folytonos függvények, akkor $f(x) > 0$ miatt elég kicsi ϵ-ra $\mathbf{P}[S_\epsilon(x)] > 0$, és így

$$\eta_i^\epsilon(x) = \frac{\mathbf{P}[Y = y_i, S_\epsilon(x)]}{\mathbf{P}[S_\epsilon(x)]} = \frac{\mathbf{P}[Y = y_i]\,\mathbf{P}[S_\epsilon(x)|Y = y_i]}{\mathbf{P}[S_\epsilon(x)]} = \frac{q_i \int_{x-\epsilon}^{x+\epsilon} f_i(z)\,dz}{\int_{x-\epsilon}^{x+\epsilon} f(z)\,dz} =$$
$$= \frac{q_i \frac{1}{2\epsilon} \int_{x-\epsilon}^{x+\epsilon} f_i(z)\,dz}{\frac{1}{2\epsilon} \int_{x-\epsilon}^{x+\epsilon} f(z)\,dz} \stackrel{\epsilon \to 0}{\longrightarrow} \frac{q_i f_i(x)}{f(x)},$$

ahol a számlálóban és a nevezőben is az integrálszámítás folytonos függvényekre vonatkozó középértéktételét alkalmaztuk. Tehát az a posteriori valószínűségek $\epsilon \to 0$ határértékben valóban a $q_i f_i(x)/f(x)$ értékek lesznek.

7. példa. Emlékezet nélküli bináris szimmetrikus csatorna (BSC) kimenetének maximum likelihood dekódolása. Tekintsük a következő szituációt: Egy emlékezet nélküli, zajos bináris szimmetrikus csatornán egy $\mathbf{Y} = (Y_1, \ldots, Y_n)$ kódszót továbbítunk. $\mathbf{Y}$ vektor valószínűségi változó, amely az n hosszúságú, bináris (például $\{0,1\}$ értékű) sorozatoknak egy rögzített $\mathcal{Y} \subseteq \{0,1\}^n$ részhalmazán (kódkönyv) veszi fel az értékét. A bináris szimmetrikus csatorna minden bitet $p \in (0,1)$ eséllyel megváltoztat, $1-p$ eséllyel változatlanul hagy. Az emlékezet nélküliség azt jelenti, hogy a különböző bitekre ezek az események függetlenek. Így a csatorna kimenete is egy n hosszúságú $\mathbf{X} = (X_1, ..., X_n) \in \mathcal{X} = \{0,1\}^n$ bitsorozat (egy vektor valószínűségi változó, amely azonban nem feltétlenül kódszó). A zajos $\mathbf{X}$-et megfigyelve akarunk döntést hozni arról, hogy mi lehetett a kódszó úgy, hogy a költségünk 1 ha hibázunk, 0 egyébként (lásd 2. példa). Figyeljük meg, hogy míg a csatorna bemenete $\mathbf{Y}$ és kimenete $\mathbf{X}$, addig a döntés (a dekódoló) bemenete $\mathbf{X}$ és kimenete $\mathbf{Y}$(-ra vonatkozó döntés).
Jelöljük az i-edik lehetséges kódszót $\mathbf{y}_i = (y_{i1}, y_{i2}, \ldots, y_{in})$-nel, egy lehetséges csatorna kimenetet $\mathbf{x} = (x_1, x_2, \ldots, x_n)$-nel, vagyis az i-edik kódszó j-edik bitje y_{ij}, a kimeneté x_j. Ha $\mathbf{Y} = \mathbf{y}_i$, akkor az $\mathbf{X} = \mathbf{x}$ feltételes valószínűsége a következő:

$$\begin{aligned}\mathbf{P}\left[\mathbf{X}{=}\mathbf{x}|\mathbf{Y}{=}\mathbf{y}_i\right] &= \mathbf{P}\left[X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n | Y_1 = y_{i1}, \ldots, Y_n = y_{in}\right] = \\ &= \mathbf{P}\left[X_1 = x_1 | Y_1 = y_{i1}\right] \mathbf{P}\left[X_2 = x_2 | Y_2 = y_{i2}\right] \ldots \mathbf{P}\left[X_n = x_n | Y_n = y_{in}\right],\end{aligned}$$

ahol az utolsó egyenlőség az emlékezet nélküliség definíciója. Mivel az átmenetvalószínűség p,

$$\mathbf{P}\left[X_t = x_t | Y_t = y_{it}\right] = \left\{ \begin{array}{ll} p, & \text{ha } x_t \neq y_{it}, \\ 1-p, & \text{ha } x_t = y_{it} \end{array} \right. = p^{\mathbf{I}_{\{x_t \neq y_{it}\}}} (1-p)^{1-\mathbf{I}_{\{x_t \neq y_{it}\}}}.$$

Így aztán

$$\begin{aligned}\mathbf{P}\left[\mathbf{X}{=}\mathbf{x}|\mathbf{Y}{=}\mathbf{y}_i\right] &= \prod_{t=1}^{n} \left[p^{\mathbf{I}_{\{x_t \neq y_{it}\}}} (1-p)^{1-\mathbf{I}_{\{x_t \neq y_{it}\}}}\right] = p^{\sum_{t=1}^n \mathbf{I}_{\{x_t \neq y_{it}\}}} (1-p)^{n-\sum_{t=1}^n \mathbf{I}_{\{x_t \neq y_{it}\}}} = \\ &= (1-p)^n \left(\frac{p}{1-p}\right)^{\sum_{t=1}^n \mathbf{I}_{\{x_t \neq y_{it}\}}}. \qquad (1.8)\end{aligned}$$

Ha az $\mathbf{Y}$ a priori eloszlása nem ismert, akkor a Bayes-döntés is ismeretlen, maximum likelihood döntést azonban használhatunk $\mathbf{X}$ ismeretében az $\mathbf{Y}$ kódszóra való következtetéshez. Ekkor azt az $\mathbf{y}_i \in \mathcal{Y}$ kódszót választjuk, amely mellett $p_i(\mathbf{x}) = \mathbf{P}\left[\mathbf{X} = \mathbf{x}|\mathbf{Y} = \mathbf{y}_i\right]$ a legnagyobb. Feltéve, hogy $0 < p < 1/2$, azaz $p/(1-p) < 1$, (1.8) alapján ez az az $\mathbf{y}_i$, amelyre $\sum_{t=1}^n \mathbf{I}_{\{x_t \neq y_{it}\}}$ a legkisebb. Megfigyelhetjük, hogy $\sum_{t=1}^n \mathbf{I}_{\{x_t \neq y_{it}\}}$ éppen azon bitek száma, amelyekben $\mathbf{x}$ és $\mathbf{y}_i$ különbözik. Ezt $\mathbf{x}$ és $\mathbf{y}_i$ *Hamming-távolságának* is nevezik. Ezek szerint bináris szimmetrikus csatorna esetén a maximum likelihood döntéssel való dekódolás a kimeneten megjelenő sorozathoz Hamming-távolságban legközelebbi kódszónak (kódszavak egyikének) a választását jelenti. (Természetesen egyenletes bemeneti ($\mathbf{Y}$) eloszlás esetén a maximum likelihood döntés most is éppen a maximum a posteriori döntés lesz.)

5. gyakorlat. a) Mi lesz a maximum likelihood dekódolás $p > 1/2$ esetén? Mi erre a magyarázat?
b) Mi lesz a maximum likelihood dekódolás $p = 1/2$ esetén? Mi erre a magyarázat?

1.4. Bayes-döntés ismételt megfigyelés alapján

Előfordul, hogy egyetlen X megfigyelés helyett n számú megfigyelés, $\mathbf{X}=(X_1,\ldots,X_n)\in\mathcal{X}^n$, is a rendelkezésünkre áll, amelyek az $\{Y=y_i\}$ feltétel mellett feltételesen függetlenek és azonos $(p_i(x))$ eloszlásúak. Tegyük fel az egyszerűség kedvéért, hogy X_t diszkrét, $k=2$ ($\mathcal{Y}=\bar{\mathcal{Y}}=\{y_1,y_2\}$) és $C_{11}=C_{22}=0$ (mint például a 2., 3. és 4. példákbeli és bármely (pszeudo)metrika tulajdonságú költségfüggvénynél). Sejthető, hogy ha az $\mathbf{X}$-ből egyáltalán lehet következtetni Y-ra, azaz az X_t-k nem függetlenek az Y-tól, akkor az $(\mathbf{X},Y)$-hoz tartozó Bayes-kockázat tetszőlegesen kicsivé válik, amint $n\to\infty$. Valóban, az alábbi tétel szerint a Bayes-kockázat exponenciálisan tart 0-hoz.

1.5. tétel. Legyen g_n^* az $(\mathbf{X},Y)$ döntési problémához tartozó Bayes-döntés. Ha $q_1q_2=0$, vagy ha $q_1q_2\neq 0$ és létezik $x\in\mathcal{X}$, hogy $p_1(x)\neq p_2(x)$, akkor

$$\lim_{n\to\infty} R(g_n^*)=0,$$

és a konvergencia exponenciálisan gyors.

4. *megjegyzés.* A tétel feltétele pontosan akkor áll fenn, ha $q_1q_2=0$, vagy ha $q_1q_2\neq 0$ és X_t és Y nem független (azaz $\sum_{i\in\{1,2\},x\in\mathcal{X}} q_ip_i(x)\log\frac{p_i(x)}{\mathbf{P}[X=x]}$ kölcsönös információjuk nem 0). Az is látható, hogy ha $q_1q_2C_{12}C_{21}\neq 0$ de $\mathbf{X}$ és Y független, akkor nem kaphatunk tetszőlegesen kicsi Bayes-kockázatot n növekedésével.

Bizonyítás. Az 1.3 fejezetben láttuk, hogy a Bayes-kockázat (1.5) szerint

$$R(g_n^*)=\mathbf{E}\left[\min(d_1(\mathbf{X}),d_2(\mathbf{X}))\right]=\sum_{\mathbf{x}\in\mathcal{X}^n}\min\left(\sum_{i=1}^2 C_{i1}\eta_i(\mathbf{x}),\sum_{i=1}^2 C_{i2}\eta_i(\mathbf{x})\right)\mathbf{P}\left[\mathbf{X}=\mathbf{x}\right]=$$
$$=\sum_{\mathbf{x}\in\mathcal{X}^n}\min(C_{21}\eta_2(\mathbf{x}),C_{12}\eta_1(\mathbf{x}))\mathbf{P}\left[\mathbf{X}=\mathbf{x}\right].$$

Ha $q_1q_2=0$, akkor $\forall\mathbf{x}\in\mathcal{X}^n$-re $\eta_1(\mathbf{x})$ vagy $\eta_2(\mathbf{x})$ is 0, így a minimum és $R(g_n^*)$ is 0. Ha $q_1q_2\neq 0$, akkor behelyettesítve (1.6)-t

$$\begin{aligned}R(g_n^*)&=\sum_{\mathbf{x}\in\mathcal{X}^n}\min\left(C_{21}\frac{q_2p_2(\mathbf{x})}{\mathbf{P}\left[\mathbf{X}=\mathbf{x}\right]},C_{12}\frac{q_1p_1(\mathbf{x})}{\mathbf{P}\left[\mathbf{X}=\mathbf{x}\right]}\right)\mathbf{P}\left[\mathbf{X}=\mathbf{x}\right]=\\&=\sum_{\mathbf{x}\in\mathcal{X}^n}\min(C_{21}q_2p_2(\mathbf{x}),C_{12}q_1p_1(\mathbf{x}))\leq\\&\leq\sum_{\mathbf{x}\in\mathcal{X}^n}\sqrt{C_{21}q_2p_2(\mathbf{x})C_{12}q_1p_1(\mathbf{x})},\end{aligned}\tag{1.9}$$

ahol a legutóbbi lépésnél azt használtuk, hogy $\forall a_1,a_2\geq 0$-ra $\min(a_1,a_2)\leq\sqrt{a_1a_2}$. $\mathbf{x}=(x_1,\ldots,x_n)\in\mathcal{X}^n$-re a $p_i(\mathbf{x})=\mathbf{P}\left[\mathbf{X}=\mathbf{x}|Y=y_i\right]$ feltételes valószínűségek az X_t-k feltételes függetlensége alapján a következőképpen fejezhetőek ki:

$$p_i(\mathbf{x})=\mathbf{P}\left[X_1=x_1,\ldots,X_n=x_n|Y=y_i\right]=\prod_{t=1}^n\mathbf{P}\left[X_t=x_t|Y=y_i\right]=\prod_{t=1}^n p_i(x_t).$$

Ezt (1.9)-be helyettesítve

$$\begin{aligned} R(g_n^*) &\le \sqrt{C_{12}C_{21}q_1q_2} \sum_{(x_1,\ldots,x_n)\in\mathcal{X}^n} \sqrt{\prod_{t=1}^n p_1(x_t) \prod_{t=1}^n p_2(x_t)} \\ &= \sqrt{C_{12}C_{21}q_1q_2} \sum_{(x_1,\ldots,x_n)\in\mathcal{X}^n} \prod_{t=1}^n \sqrt{p_1(x_t)p_2(x_t)} \\ &= \sqrt{C_{12}C_{21}q_1q_2} \left(\sum_{x\in\mathcal{X}} \sqrt{p_1(x)p_2(x)} \right)^n . \end{aligned}$$

A számtani és mértani közép közötti összefüggés alapján

$$\sum_{x\in\mathcal{X}} \sqrt{p_1(x)p_2(x)} \le \sum_{x\in\mathcal{X}} \frac{p_1(x)+p_2(x)}{2} = \frac{1}{2}\sum_{x\in\mathcal{X}} p_1(x) + \frac{1}{2}\sum_{x\in\mathcal{X}} p_2(x) = 1,$$

ahol egyenlőség (akkor és) csak akkor áll fenn, ha $\forall x \in \mathcal{X}$-re $p_1(x) = p_2(x)$. Tehát ha $\exists x \in \mathcal{X}$, hogy $p_1(x) \neq p_2(x)$, akkor $0 \le \sum_{x\in\mathcal{X}} \sqrt{p_1(x)p_2(x)} < 1$, így $R(g_n^*)$ exponenciálisan 0-hoz tart. □

1.5. Bayes-döntés közelítése

A $d_i(x)$ várható költségek (vagy az $\eta_i(x)$ a posteriori valószínűségek) pontos értékei általában ismeretlenek. Tekintsünk egy ilyen szituációt, amelyben azonban a d_i-ket meg tudjuk becsülni valamely $\tilde{d}_i$ függvényekkel. Az (1.4)-et teljesítő Bayes-döntés mintájára a $\{\tilde{d}_i\}_{1\le i\le k}$ függvényekhez is hozzárendelhetünk analóg módon olyan $\tilde{g}$ döntésfüggvényt, amely x esetén az (egyik) legkisebb $\tilde{d}_j(x)$-hez tartozó y_j-re dönt, azaz $\tilde{g}(x) = y_j \Rightarrow \tilde{d}_j(x) = \min_{1\le i\le k} \tilde{d}_i(x)$. (Több minimális esetén valamilyen elv szerint választunk ezen hipotézisek közül, például ismét a legkisebb indexűt.) $\tilde{g}$ tehát úgy viszonyul a $\tilde{d}_i$-khez, ahogy g^* a d_i-khez. Vajon ha a $\tilde{d}_i$-k jó becslések, akkor $\tilde{g}$ (lokális/globális) kockázata közel lesz g^* (lokális/globális) kockázatához? (Kisebb természetesen nem lehet a 1.4. tétel alapján.) A következő tétel erre ad pozitív választ, megmutatva, hogy a szóbanforgó kockázatok különbsége korlátozható a $\tilde{d}_i$-k becslési hibájával:

1.6. tétel. $i = 1,\ldots,k$-ra legyen $\tilde{d}_i : \mathcal{X} \to [0,\infty)$ a d_i becslése és $\tilde{g}$ egy a $\{\tilde{d}_i\}_{1\le i\le k}$-khez rendelt döntésfüggvény. Ekkor

$$r(\tilde{g},x) - r(g^*,x) \le \mathbf{I}_{\{\tilde{g}(x)\neq g^*(x)\}} \sum_{i=1}^k |\tilde{d}_i(x) - d_i(x)|$$

és

$$R(\tilde{g}) - R(g^*) \le \mathbf{E}\left[\mathbf{I}_{\{\tilde{g}(X)\neq g^*(X)\}} \sum_{i=1}^k |\tilde{d}_i(X) - d_i(X)| \right] \le \mathbf{E}\left[\sum_{i=1}^k |\tilde{d}_i(X) - d_i(X)| \right].$$

Bizonyítás. Az egyszerűbb jelölés kedvéért az általánosság megszorítása nélkül feltehetjük, hogy $y_j = j$. Ekkor (1.3) szerint a lokális kockázata így is írható:

$$r(g,x) = \sum_{j=1}^{k} \mathbf{I}_{\{g(x)=j\}} d_j(x) = d_{g(x)}(x).$$

Tehát $\tilde{g}$ és g^* lokális kockázatának különbsége

$$r(\tilde{g},x) - r(g^*,x) = d_{\tilde{g}(x)}(x) - d_{g^*(x)}(x).$$

Ha $\tilde{g}(x) = g^*(x)$, akkor $r(\tilde{g},x) - r(g^*,x) = 0$. Ha $\tilde{g}(x) \neq g^*(x)$, akkor viszont

$$\begin{aligned}
d_{\tilde{g}(x)}(x) - d_{g^*(x)}(x) &= d_{\tilde{g}(x)}(x) - \tilde{d}_{\tilde{g}(x)}(x) + \tilde{d}_{\tilde{g}(x)}(x) - d_{g^*(x)}(x) \\
&\leq d_{\tilde{g}(x)}(x) - \tilde{d}_{\tilde{g}(x)}(x) + \tilde{d}_{g^*(x)}(x) - d_{g^*(x)}(x) \\
&\Big(\text{mert } \tilde{g} \text{ definíciója szerint } \tilde{d}_{\tilde{g}(x)}(x) \leq \tilde{d}_{g^*(x)}(x)\Big) \\
&\leq |d_{\tilde{g}(x)}(x) - \tilde{d}_{\tilde{g}(x)}(x)| + |\tilde{d}_{g^*(x)}(x) - d_{g^*(x)}(x)| \\
&\leq \sum_{i=1}^{k} |\tilde{d}_i(x) - d_i(x)| \\
&\quad (\text{mert } \tilde{g}(x) \text{ és } g^*(x) \text{ különböző elemei } \{1,\ldots,k\}\text{-nak}).
\end{aligned}$$

Összefoglalva:

$$r(\tilde{g},x) - r(g^*,x) \leq \mathbf{I}_{\{\tilde{g}(x)\neq g^*(x)\}} \sum_{i=1}^{k} |\tilde{d}_i(x) - d_i(x)|,$$

majd x helyére X-et írva és várhatóértéket véve kapjuk a korlátot $R(\tilde{g}) - R(g^*)$-re. Az utolsó egyenlőtlenség triviális $\mathbf{I}_{\{\tilde{g}(x)\neq g^*(x)\}} \leq 1$-ből. □

Nézzük ismét a 2. példabeli $C_{ij} = \mathbf{I}_{\{i\neq j\}}$ speciális esetet, amikor – mint láttuk – $d_i(x) = 1 - \eta_i(x)$ és $g^*(x) = y_j \Rightarrow \eta_j(x) = \max_{1\leq i\leq k} \eta_i(x)$. Feltehető, hogy ekkor a $\tilde{d}_i$ becslők $1 - \tilde{\eta}_i$ alakban állnak elő, ahol $\tilde{\eta}_i$-k a η_i-k becslései. Ekkor

$$\tilde{g}(x) = y_j \Rightarrow \tilde{\eta}_j(x) = \max_{1\leq i\leq k} \tilde{\eta}_i(x), \tag{1.10}$$

és a 1.6. tétel alakja

$$r(\tilde{g},x) - r(g^*,x) \leq \mathbf{I}_{\{\tilde{g}(x)\neq g^*(x)\}} \sum_{i=1}^{k} |\tilde{\eta}_i(x) - \eta_i(x)|$$

és

$$R(\tilde{g}) - R(g^*) \leq \mathbf{E}\left[\mathbf{I}_{\{\tilde{g}(X)\neq g^*(X)\}} \sum_{i=1}^{k} |\tilde{\eta}_i(X) - \eta_i(X)|\right] \leq \mathbf{E}\left[\sum_{i=1}^{k} |\tilde{\eta}_i(X) - \eta_i(X)|\right] \tag{1.11}$$

lesz. Ha $k = 2$, és feltesszük, hogy $(\tilde{\eta}_1(x), \tilde{\eta}_2(x))$ minden $x \in \mathcal{X}$-re eloszlást alkot, azaz $\tilde{\eta}_2(x) = 1 - \tilde{\eta}_1(x)$, akkor ezek tovább egyszerűsödnek a következőképpen:

$$r(\tilde{g}, x) - r(g^*, x) \leq \mathbf{I}_{\{\tilde{g}(x) \neq g^*(x)\}} \left(|\tilde{\eta}_1(x) - \eta_1(x)| + |\tilde{\eta}_2(x) - \eta_2(x)|\right) = \\ = 2\mathbf{I}_{\{\tilde{g}(x) \neq g^*(x)\}} |\tilde{\eta}_1(x) - \eta_1(x)|$$

és

$$R(\tilde{g}) - R(g^*) \leq 2\mathbf{E}\left[\mathbf{I}_{\{\tilde{g}(X) \neq g^*(X)\}} |\tilde{\eta}_1(X) - \eta_1(X)|\right] \leq 2\mathbf{E}\left[|\tilde{\eta}_1(X) - \eta_1(X)|\right].$$

8. példa. Legyen $C_{ij} = \mathbf{I}_{\{i \neq j\}}$, $k = 2$, $y_1 = 1$ és $y_2 = 0$. Mutassuk meg, hogy ekkor $\eta_1(X) = \mathbf{E}[Y|X]$. Tehát ha van egy jó becslésünk Y feltételes várhatóérték függvényére, akkor az ahhoz (1.10) alapján rendelt döntésfüggvény közel optimális.

1.6. Bayes-becslés

Szemben az eddigi 1.3–1.5. fejezettel, ahol $\mathcal{Y} = \bar{\mathcal{Y}}$ (tetszőleges) véges halmaz volt, és többnyire a 0-1 költséget vizsgáltuk, azaz Y értékét el akartuk *dönteni* és hiba esetén nem volt érdekes, hogy mekkora a hibázás nagysága, a következő fejezetekben $\mathcal{Y}$ és $\bar{\mathcal{Y}}$ általában $\mathbf{R}$ részhalmaza, többnyire végtelen, sőt folytonos, és C valamiféle távolságát méri a valódi és a *becsült* paraméternek, tehát a hibázás nagysága is számít. Ekkor g-t *becslésfüggvénynek*, $g(X)$-t *becslésnek* is nevezzük. A Bayes-feladat továbbra is az $R(g)$ globális kockázatot minimalizáló $g : \mathcal{X} \to \mathcal{Y}$ becslésfüggvény megtalálása. (Például a 4. példabeli C_2-re $R(g)$ éppen a négyzetes középhiba.)

5. definíció. *Bayes-becslésnek* nevezzük az optimális g^* becslésfüggvényeket, azaz amelyekre

$$R^* \stackrel{\text{def}}{=} R(g^*) = \min_g R(g).$$

Fennáll a következő elégséges feltétel:

1.7. tétel. Ha egy g^* becslésfüggvény lokális kockázatára

$$r(g^*, x) = \min_{y \in \bar{\mathcal{Y}}} \mathbf{E}\left[C(Y, y)|X = x\right], \qquad \forall x \in \mathcal{X},$$

akkor g^* Bayes-becslés.

Bizonyítás. Bármely g becslésfüggvényre

$$R(g) = \mathbf{E}\left[\mathbf{E}\left[C(Y, g(X))|X\right]\right] \geq \mathbf{E}\left[\min_{y \in \bar{\mathcal{Y}}} \mathbf{E}\left[C(Y, y)|X\right]\right] = \mathbf{E}\left[r(g^*, X)\right] = R(g^*),$$

így g^* optimális. □

9. példa. Ha $\mathcal{Y}$ ismét véges, és tekintjük a 2. példabeli $C_0(y, y') = \mathbf{I}_{\{y \neq y'\}}$ költséget, akkor nyilván nincs ok olyan g becslésfüggvényt használni, amelynek értékkészlete nem része $\mathcal{Y}$-nak. Így ekkor a becslési feladat ekvivalens lesz a 2. példabeli döntési feladattal, és a Bayes-becslést az 1.3 fejezet szerinti *maximum a posteriori becslés* (döntés) adja, azaz ha $g^*(x) = y_j$, akkor $\eta_j(x) \geq \eta_i(x)$ $\forall i \neq j$-re (vagyis $\eta_j(x) = \max_{1 \leq i \leq k} \eta_i(x)$), ahol az $\eta_i(x)$-eket (1.6), illetve (1.7)-ből számolhatjuk ki, ha X diszkrét, illetve abszolút folytonos. Tehát a maximális a posteriori becslést – és így a Bayes-becslést – az a hipotézis adja, amelyikre $q_i p_i(x)$, illetve $q_i f_i(x)$ maximális.

1.7. Maximum likelihood becslés

Mivel most Y tipikusan nem diszkrét, hanem például abszolút folytonos, ez esetben Y q_i a priori valószínűségei helyett csak a $q(y)$ a priori sűrűségfüggvénye értelmezhető. Ekkor X feltételes ($Y = y_i$ melletti) $p_i(x)$ súlyfüggvényei, illetve $f_i(x)$ sűrűségfüggvényei helyett a

$$p_y(x) = \mathbf{P}\left[X = x | Y = y\right]$$

feltételes súlyfüggvényei, illetve

$$f_y(x) = f(x|Y = y)$$

feltételes sűrűségfüggvényei értelmezhetők, amelyek az X (diszkrét, illetve abszolút folytonos) eloszlását adják meg $Y = y \in \mathcal{Y}$ feltétel esetén. Ugyan formálisan ekkor is definiálható lenne az $\arg\max_{y\in\mathcal{Y}} q(y)p_y(x)$, illetve $\arg\max_{y\in\mathcal{Y}} q(y)f_y(x)$ maximum a posteriori becslés, ez nem releváns, mert folytonos Y-ra $\forall y \in \bar{\mathcal{Y}}$-ra $\mathbf{E}\left[C_0(Y, y)\right] = \mathbf{P}\left[Y \neq y\right] \equiv 1$ 1 valószínűséggel.
A *maximum likelihood becslés* azonban definiálható (a maximum likelihood döntéssel analóg módon):

6. definíció. g-t maximum likelihood becslésnek nevezzük, ha $p_{g(x)}(x) = \max_{y\in\mathcal{Y}} p_y(x)$, illetve $f_{g(x)}(x) = \max_{y\in\mathcal{Y}} f_y(x)$ diszkrét, illetve abszolút folytonos X esetén.

10. példa. Az 1.4. fejezetben tárgyaltakhoz hasonlóan legyen az $\mathbf{X} = (X_1, \ldots, X_n) \in \mathbf{R}^n$ vektor valószínűségi változó, ahol az X_t-k független, ismeretlen Y várhatóértékű, σ szórású, azonos normális (Gauss-) eloszlású valószínűségi változók. Határozzuk meg Y-nak az $\mathbf{X}$ megfigyelésre alapozott maximum likelihood becslését!
$\mathbf{X}$ feltételes sűrűségfüggvényét $Y = y$ esetén ekkor a többdimenziós normális sűrűségfüggvény adja:

$$f_y(\mathbf{x}) = \frac{1}{\left(\sqrt{2\pi}\sigma\right)^n} e^{-\frac{1}{2\sigma^2}\sum_{t=1}^n (x_t - y)^2},$$

ahol $\mathbf{x} = (x_1, \ldots, x_n)$. A maximum likelihood becslés az az y lesz, amelyre $f_y(\mathbf{x})$ maximális. Ez megegyezik azzal, amelyre

$$-\ln f_y(\mathbf{x}) = n\ln(\sqrt{2\pi}\sigma) + \frac{1}{2\sigma^2}\sum_{t=1}^n (x_t - y)^2,$$

azaz $\frac{1}{n}\sum_{t=1}^n (x_t - y)^2$ minimális. Legyen $m_n \stackrel{\text{def}}{=} \frac{1}{n}\sum_{t=1}^n x_t$. Ekkor a 1.1. tételt alkalmazva egy, az $(x_1, \ldots, x_n)$-en egyenletes – és így m_n várhatóértékű – valószínűségi változóra:

$$\frac{1}{n}\sum_{t=1}^n (x_t - y)^2 = \frac{1}{n}\sum_{t=1}^n (x_t - m_n)^2 + (m_n - y)^2,$$

amely nyilván $y = m_n$-re minimális. Így a várhatóérték maximum likelihood becslése normális $\mathbf{X}$-re éppen a megfigyelések átlaga.

11. példa. A fentiekhez hasonlóan legyen most az $\mathbf{X} = (X_1, ..., X_n) \in \mathbf{R}^n$, ahol az X_i-k független, ismeretlen Y várhatóértékű, Σ szórású, azonos, de egyenletes eloszlású valószínűségi változók. Határozzuk meg az (Y, Σ) párnak az $\mathbf{X}$ megfigyelésre alapozott maximum likelihood becslését!
$Y = y$, $\Sigma = \sigma$ esetén minden X_t-nek az $[y - \sqrt{3}\sigma, y + \sqrt{3}\sigma]$ intervallumon kell egyenletesnek lennie (a szórásnégyzet ekkor lesz $(2\sqrt{3}\sigma)^2/12 = \sigma^2$), így – bevezetve $\delta \stackrel{\text{def}}{=} \sqrt{3}\sigma$-t – feltételes sűrűségfüggvényük:

$$g_{y,\sigma} = \frac{\mathbf{I}_{[y-\delta, y+\delta]}}{2\delta}.$$

Tehát $\mathbf{X}$ feltételes sűrűségfüggvénye, ha $Y = y$, $\Sigma = \sigma$ a függetlenség miatt $f_{y,\sigma}(\mathbf{x}) = \prod_{t=1}^{n} g_{y,\sigma}(x_t)$. Így ez ugyanarra az (y, σ) párra veszi fel maximumát, amelyre

$$\ln f_{y,\sigma}(\mathbf{x}) = \sum_{t=1}^{n} \ln g_{y,\sigma}(x_t) = \sum_{t=1}^{n} \ln \frac{\mathbf{I}_{\{x_t \in [y-\delta, y+\delta]\}}}{2\delta} = \sum_{t=1}^{n} \ln \mathbf{I}_{\{x_t \in [y-\delta, y+\delta]\}} - n \ln(2\delta) =$$
$$= \begin{cases} -n \ln(2\delta), & \text{ha } \forall x_t \in [y - \delta, y + \delta], \\ -\infty, & \text{egyébként.} \end{cases} \tag{1.12}$$

Ez akkor maximális, ha minden $x_t \in [y - \delta, y + \delta]$ a lehető legkisebb δ mellett. Az $[y - \delta, y + \delta]$ intervallumnak tehát a legkisebb x_t-től a legnagyobbig kell tartania, azaz az (Y, Σ) maximum likelihood becslése

$$y = \frac{\underline{X} + \overline{X}}{2}, \qquad \text{illetve} \qquad \sigma = \frac{\overline{X} - \underline{X}}{2\sqrt{3}},$$

ahol $\underline{X} = \min_{1 \le t \le n} X_t$ és $\overline{X} = \max_{1 \le t \le n} X_t$.

12. példa. Tekintsük a 11. példabeli helyzetet, de legyen most a σ szórás – tehát az intervallumhossz – ismert. Határozzuk meg az Y-nak az $\mathbf{X}$ megfigyelésre alapozott maximum likelihood becslését!
A számolás nem változik, most is a (1.12)-t kell maximalizálni, de csak y-ban, míg $\delta = \sqrt{3}\sigma$ ismert, rögzített paraméter. (1.12) akkor maximális, ha minden $X_t \in [y-\delta, y+\delta]$, azaz $\overline{X} - \delta \le y \le \underline{X} + \delta$. Minden ilyen y az Y egy maximum likelihood becslése. Ez mutatja, hogy a maximum likelihood becslés nem mindig egyértelmű.

5. *megjegyzés.* Vegyük észre, hogy a 11. és a 12. példánál is az $(\underline{X}, \overline{X})$ pár ismerete elégséges volt a maximum likelihood becslés meghatározásához.

1.8. Regresszióbecslés; négyzetes középhiba minimalizálás

Legyen $Y \in \mathbf{R}$ véges szórású változó, és vizsgáljuk a 4. példabeli $C_2(y, y') = (y - y')^2$ négyzetes költséget, azaz keressük azt a g^* becslésfüggvényt, amelyre $R(g^*) = \mathbf{E}\left[(g^*(X) - Y)^2\right]$ minimális. A 1.7. tétel szerint ha

$$r(g^*, x) = \min_{y \in \mathcal{Y}} \mathbf{E}\left[(Y - y)^2 | X = x\right], \qquad \forall x \in \mathcal{X},$$

akkor g^* Bayes-becslés. Definiáljuk a *regressziós függvény* fogalmát:

7. definíció. Regressziós függvénynek nevezzük az

$$m(x) = \mathbf{E}\left[Y|X = x\right]$$

függvényt, amely minden x-re Y-nak a feltételes várhatóértékét adja, ha $X = x$.

1.8. tétel. Négyzetes költség esetén $g^*(X) = m(X)$ 1 valószínűséggel, vagyis a Bayes-becslés éppen a regressziós függvény. [3, p. 2]

Bizonyítás. Rögzített x-re alkalmazzuk a 1.1. tételt az $Y|X = x$ feltételes eloszlásra: bármely $g(x) \in \mathbf{R}$-re

$$\begin{aligned} r(g,x) &= \mathbf{E}\left[(Y - g(x))^2|X = x\right] = \\ &= \mathbf{E}\left[(Y - \mathbf{E}\left[Y|X = x\right])^2|X = x\right] + (\mathbf{E}\left[Y|X = x\right] - g(x))^2 = \\ &= \mathbf{E}\left[(Y - m(x))^2|X = x\right] + (m(x) - g(x))^2 = r(m,x) + (m(x) - g(x))^2. \end{aligned}$$

Tehát bármely g becslésre

$$R(g) = \mathbf{E}\left[r(g,X)\right] = \mathbf{E}\left[r(m,X)\right] + \mathbf{E}\left[(m(X) - g(X))^2\right] = R(m) + \mathbf{E}\left[(m(X) - g(X))^2\right],$$

ami pontosan akkor minimális, ha $g(X) = m(X)$ 1 valószínűséggel. □

6. *megjegyzés.* A 1.8. tétel szerint ha Y-t négyzetes értelemben akarjuk közelíteni az X egy $g(X)$ függvényével, akkor az $\mathbf{E}\left[(g(X) - Y)^2\right]$ középhibát minimalizáló $g(x)$ az $\mathbf{E}\left[Y|X = x\right]$ feltételes várhatóérték. Azaz csupán x ismeretében ez a legjobb becslés.

6. gyakorlat. Bizonyítsuk be, hogy ha $(X,Y) \in \mathbf{R}^2$ együttesen normális eloszlású és – az egyszerűség kedvéért – mindkettő szórása 1, akkor a regressziós függvény

$$m(x) = \mathbf{E}\left[Y|X = x\right] = \mathbf{E}\left[Y\right] + \rho(x - \mathbf{E}\left[X\right]),$$

ahol $\rho = \mathbf{E}\left[XY\right] - \mathbf{E}\left[X\right]\mathbf{E}\left[Y\right]$ az X és Y korrelációs együtthatója (egyben kovarianciája). Próbáljuk meg általánosítani ezt tetszőleges szórásokra, majd többdimenziós $\mathbf{X}$ vektor változóra.

1.8.1. Lineáris becslés

Továbbra is legyen $Y \in \mathbf{R}$ véges szórású, $\mathbf{X} = (X_1, ..., X_d) \in \mathbf{R}^d$ pedig vektor valószínűségi változó, és vizsgáljuk a $C_2(y,y') = (y - y')^2$ költséget. A 1.8. tétel szerint a Bayes-becslést lényegében az $m(x)$ regressziós függvény adja. A 6. gyakorlatban láthattuk, hogy ha $(\mathbf{X},Y)$ együttesen normális eloszlású (az egyéb speciális tulajdonságai miatt ez az egyik legfontosabb modell), akkor az $m(\mathbf{x})$ regressziós függvény az $\mathbf{x}$ megfigyelés lineáris függvénye. Ebben az esetben tehát elég a Bayes-becslést $\mathbf{x}$ lineáris függvényei között keresni. A lineáris függvények igen tömören tárolhatóak és könnyen kiértékelhetőek. Bár a Bayes-becslés általános (nem gaussi) esetben nem mindig lineáris függvény, a fentiek miatt érdemes ekkor is alkalmazni azt a megszorítást, hogy a becslésfüggvényt csak $\mathbf{x}$-nek a lineáris függvényei között keressük. Természetesen a

$$g(\mathbf{x}) = c_0 + \sum_{i=1}^{d} c_i x_i \qquad c_0, c_1, \ldots, c_d \in \mathbf{R}$$

lineáris becslésfüggvények közül a legkisebb kockázatút szeretnénk megtalálni. Az egyszerűbb jelölés kedvéért vezessünk be az $\mathbf{X}$ (illetve $\mathbf{x}$) vektor egy nulladik X_0 (illetve x_0) koordinátáját, amely azonosan egyenlő 1-gyel ($X_0 \equiv x_0 \equiv 1$). Az alábbiakban tehát $\mathbf{X}$,$\mathbf{x} \in \mathbf{R}^{d+1}$. Így $g(\mathbf{x}) = \sum_{i=0}^{d} c_i x_i$, és g globális kockázata

$$R(g) = \mathbf{E}\left[C_2(Y, g(\mathbf{X}))\right] = \mathbf{E}\left[\left(Y - \sum_{i=0}^{d} c_i X_i\right)^2\right] \stackrel{\text{def}}{=} R(c_0, c_1, \ldots, c_d).$$

Keressük tehát azokat a $(c_0^*, c_1^*, \ldots, c_d^*)$ együtthatókat, amelyekre

$$\bar{g}(\mathbf{x}) = \sum_{i=0}^{d} c_i^* x_i$$

a legkisebb négyzetes költséget adja, azaz amelyekre

$$R(c_0^*, c_1^*, \ldots, c_d^*) = \min_{(c_0, c_1, \ldots, c_d) \in \mathbf{R}^{d+1}} R(c_0, c_1, \ldots, c_d),$$

azaz minimalizálni akarjuk $R(c_0, c_1, \ldots, c_d)$-t a c_i együtthatók szerint. A $(c_0^*, c_1^*, \ldots, c_d^*)$ minimumhelyen a $R(c_0, c_1, \ldots, c_d)$-nek minden változója szerinti parciális deriváltja 0 kell legyen. Mivel $\mathbf{E}[cZ] = c\mathbf{E}[Z]$ bármely c konstansra és Z valószínűségi változóra, a várhatóérték lineáris operátor, aminek alapján belátható, hogy a deriválással felcserélhető. Így

$$\frac{\partial}{\partial c_j} R(c_0, c_1, \ldots, c_d) = \mathbf{E}\left[\frac{\partial}{\partial c_j}\left(Y - \sum_{i=0}^{d} c_i X_i\right)^2\right] = \mathbf{E}\left[2\left(Y - \sum_{i=0}^{d} c_i X_i\right)(-X_j)\right] =$$

$$= 2\left(\sum_{i=0}^{d} c_i \mathbf{E}[X_i X_j] - \mathbf{E}[X_j Y]\right) \stackrel{\text{def}}{=} 2\left(\sum_{i=0}^{d} c_i s_{ij} - b_j\right),$$

ahol $s_{ij} = \mathbf{E}[X_i X_j]$, $b_j = \mathbf{E}[X_j Y]$ (i,$j = 0$,1,…,d). Következésképpen

$$\sum_{i=0}^{d} c_i^* s_{ij} = b_j, \qquad j = 0, 1, \ldots, d.$$

Térjünk át mátrixos jelölésre; legyen

$$\mathbf{b}^T = (b_0, b_1, \ldots, b_d) \qquad \text{(sorvektor)},$$
$$\mathbf{c}^{*T} = (c_0^*, c_1^*, \ldots, c_d^*) \qquad \text{(sorvektor)},$$
$$\mathbf{S} = [s_{ij}] \qquad ((d+1) \times (d+1)\text{-es mátrix}).$$

(Ha az $\mathbf{X}$ oszlopvektort jelöl, transzponáltját pedig $\mathbf{X}^T$, akkor $\mathbf{S} = \mathbf{E}\left[\mathbf{X}\mathbf{X}^T\right]$ alakban is írható.) Ezekkel a fenti lineáris egyenletrendszer a

$$\mathbf{c}^{*T}\mathbf{S} = \mathbf{b}^T$$

mátrixegyenletként írható fel. Ha tehát az $\mathbf{S}$ invertálható, akkor az egyenletrendszer egyértelmű megoldása

$$\mathbf{c}^{*T} = \mathbf{b}^T\mathbf{S}^{-1}.$$

Ezek az együtthatók adják tehát az optimális lineáris regresszióbecslést az $\{s_{ij}\}$ és $\{b_j\}$ várhatóértékek függvényében.

7. *megjegyzés.* Ha a megfigyelés centrált, azaz $\mathbf{E}[X_i] = 0$ $(i = 1, \ldots, d)$, akkor $\mathbf{S}$-nek az $(X_1, \ldots, X_d)$-hez tartozó főminorja éppen e vektornak a kovarianciamátrixa, ($\mathbf{S}$ pedig $\mathbf{X}$ kovarianciamátrixa azzal az eltéréssel, hogy $s_{00} = 1$, míg a kovarianciamátrixban itt 0 áll.)

8. *megjegyzés.* Vegyük észre, hogy $\mathbf{S}$ mindig pozitív szemidefinit, azaz minden $\mathbf{c} \in \mathbf{R}^{d+1}$-re $\mathbf{c}^T\mathbf{S}\mathbf{c} \geq 0$. Ugyanis

$$\sum_{i,j=0}^{d} c_i s_{ij} c_j = \sum_{i,j=0}^{d} c_i \mathbf{E}[X_i X_j] c_j = \mathbf{E}\left[\left(\sum_{i=0}^{d} c_i X_i\right)\left(\sum_{j=0}^{d} c_j X_j\right)\right] = \mathbf{E}\left[\left(\sum_{i=0}^{d} c_i X_i\right)^2\right] \geq 0.$$

Ismeretes, hogy egy pozitív szemidefinit mátrix pontosan akkor invertálható, ha pozitív definit, azaz ha a fenti egyenlőtlenségben nagyobb-egyenlőség helyett szigorú $>$ áll fenn minden $\mathbf{c} \in \mathbf{R}^{d+1} \setminus \{\mathbf{0}\}$-ra. Könnyű belátni, hogy ez akkor van így, ha a $(X_1, \ldots, X_d)$ vektor nem koncentrálódik $\mathbf{R}^d$ egyetlen d-nél kisebb dimenziós eltolt alterére sem.

Irodalomjegyzék

[1] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* Wiley, New York, NY, 1991.

[2] L. Devroye, L. Györfi, and G. Lugosi. *A Probabilistic Theory of Pattern Recognition.* Springer-Verlag, New York, NY, 1996.

[3] L. Györfi, M. Kohler, A. Krzyzak, and H. Walk. *A Distribution-Free Theory of Nonparametric Regression.* Springer, New York, NY, 2002.

[4] T. Linder and G. Lugosi. *Bevezetés az Információelméletbe.* Tankönyvkiadó, Budapest, 1990. jegyzetszám: J5-1445.

Jelölések*

Felhasznált jelölések

x,$\underline{x}$,$\underline{\underline{x}}$	skalár, (oszlop)vektor vagy halmaz, mátrix
$X, x, p(X)$	véletlen változó X, érték x, valószínűségi tömegfüggvény/sűrűségfüggvény X
$\mathrm{E}_{X,p(X)}[f(X)]$	$f(X)$ várható értéke $p(X)$ szerint
$\mathrm{var}_{p(X)}[f(X)]$	$f(X)$ varianciája $p(X)$ szerint
$I_p(\underline{X}\|\underline{Z}\|\underline{Y})$	$\underline{X}$ és $\underline{Y}$ megfigyelési függetlensége $\underline{Z}$ feltétellel p esetében
$(X \perp\!\!\!\perp Y\|Z)_p$	$I_p(\underline{X}\|\underline{Z}\|\underline{Y})$
$(X \not\perp\!\!\!\perp Y\|Z)_p)$	$\neg I_p(\underline{X}\|\underline{Z}\|\underline{Y})$
$CI_p(\underline{X};\underline{Y}\|\underline{Z})$	$\underline{X}$ és $\underline{Y}$ beavatkozási függetlensége $\underline{Z}$ feltétellel p esetében
$\prec$	(részleges) sorrendezés
$\prec^c$	a változók egy teljes sorrendezése
$\prec^G$	adott G irányított körmentes gráffal kompatibilis sorrendek halmaza
$\prec(n)$	n objektum sorrendjeinek (permutációinak) a halmaza
$G, \underline{\theta}$	Bayes-háló struktúrája és paraméterei
$G^{\sim}$	G irányított körmentes gráf esszenciális gráfja
$\mathcal{G}(n)/\mathcal{G}^k(n)$	n csomópontú maximum k szülőjű DAG-ok halmaza
$\mathcal{G}^{\prec}$	adott $\prec$ sorrenddel kompatibilis DAG-ok halmaza
$\mathcal{G}^G$	adott G DAG-gal megfigyelési ekvivalens DAG-ok halmaza
$\sim$	kompatibilitási reláció
$\mathrm{pa}(X_i, G) \sim \prec$	$\mathrm{pa}(X_i, G)$ szülői halmaz kompatibilis $\prec$ sorrendezéssel
$\mathrm{MB}_p(X_i)$	Markov-takarója X_i-nek p-ben
$\mathrm{pa}, \mathrm{pa}(X_i, G)$	szülői változók halmaza, X_i szüleinek halmaza G-ben
pa_{ij}	a j. konfigurációja a szülői értékeknek egy sorrendben
$\mathrm{bd}(X_i, G)$	X_i szüleinek, gyerekeinek és gyerekei egyéb szüleinek halmaza G-ben
$\mathrm{MBG}(X_i, G)$	a Markov-takaró algráfja X_i-nek G-ben
$\mathrm{MBM}(X_i, X_j, G)$	a Markov-takaróbeliség relációja
n	valószínűségi változók száma
k	maximális szülőszám DAG-okban
N	mintaszám
V	összes valószínűségi változók száma
Y	válasz, kimeneteli, függő változó
$N_+/N_{\ldots,+,\ldots}$	$N_i/N_{\ldots,i,\ldots}$ megfelelő összegei
$D\|X$	X változóhalmazra szűkített adathalmaz
$\|\|$	kardinalitás

*További konvenciók az egyes fejezetekben jelöltek.

$1()$	indikátorfüggvény
f', f''	f függvény első és második deriváltjai
A^T	A mátrix transzponáltja
$\underline{x} \cdot \underline{y}$	$\underline{x}$ és $\underline{y}$ vektorok skalárszorzata
ξ^+/ξ^-	informatív/nem informatív információs kontextus
$\neg, \wedge, \vee, \not\equiv, \rightarrow$	standard logikai operátorok
$\cap, \cup, \setminus, \Delta$	standard halmazműveletek
$KB \vdash_i \alpha$	α bizonyíthatósága KB-ből
Γ	a Gamma függvény
$\mathrm{Beta}(x\|\alpha, \beta)$	a Béta eloszlás sűrűségfüggvénye (pdf)
$\mathrm{Dir}(x\|\underline{\alpha})$	a Dirichlet eloszlás sűrűségfüggvénye
$\mathrm{N}(x\|\mu, \sigma)$	az egyváltozós normál eloszlás sűrűségfüggvénye
$\mathrm{N}(x\|\underline{\mu}, \underline{\underline{\Sigma}})$	a többváltozós normál eloszlás sűrűségfüggvénye
BD,BD_e	Bayesian Dirichlet prior, megfigyelési ekvivalens BD prior
BD_{CH}	Bayesian Dirichlet (BD) prior 1 hiperparaméterekkel
$\mathrm{BD}_{\mathrm{eu}}$	megfigyelési ekvivalens és uniform BD prior
$L(\underline{\theta}; D_N)$	$p(D_N\|\underline{\theta})$ likelihood függvénye
$H(X, Y)$	X és Y entrópiája
$I(X; Y)$	X és Y kölcsönös információja
$\mathrm{KL}(X\|Y)$	X és Y Kullback–Leibler divergenciája
$H(X\|Y)$	X és Y keresztentrópiája
$L_1(,)$, $L_2(,)$	az abszolútértékbeli (Manhattan) négyzetes (euklidészi) távolságok
$L_0(,)$	0-1 veszteség
$\mathcal{O}()/\Theta()$	aszimptotikus, nagyságrendi felső és alsó határ

Rövidítések

ROC	Receiver Operating Characteristic (ROC) görbe
AUC	ROC-görbe alatti terület
BMA	bayesi modell átlagolás
BN	Bayes-háló
DAG	irányított körmentes gráf
FSS	jegykiválasztási probléma
MAP	maximum a posteriori
MI	kölcsönös információ
ML	maximum likelihood
MBG	Markov-határ gráf
MB	Markov-takaró
MBM	Markov-takaróbeliség
(MC)MC	(Markov-láncos) Monte Carlo
NBN	naiv Bayes-háló

2. fejezet

Valószínűségi gráfos modellek

A fejezetben összefoglaljuk a Bayes-statisztikai keretet, különösen a megközelítés által igényelt bizonytalanságokat kezelő modell tulajdonságait. Ezt követően a valószínűségi eloszlások strukturális jellemzőit vizsgáljuk meg, ami a strukturális jellemzők explicit reprezentálásán át elvezet a valószínűségi gráfos modellosztályok használatához. A valószínűségi gráfos modellosztályon belül elsőként az egyszerű Naív Bayes-háló, Markov-lánc és rejtett Markov modell modelltípusokat foglaljuk össze, majd a Bayes-hálók és a Markov-hálók általános definícióit adjuk meg és tekintjük át. Megvizsgáljuk a reprezentációk pontosságát, kiterjesztéseiket és azok tudásmérnöki vonatkozásait.

2.1. Bevezetés

A jegyzetben a bizonytalanság különböző megjelenési formáit kizárólagosan a valószínűségelmélet keretében formalizáljuk. A valószínűségi alapokon történő megközelítés mellett, különösen a szubjektív valószínűségi értelmezéshez szorosan kapcsolódó bayesi megközelítés mellett számos érvet fel lehet sorakoztatni. A terjedelmi korlátokon belül ezek közül többet is bemutatunk, amelyek közül kiemelkednek az axiomatikus (avagy meta-axiomatikusnak is nevezhető) megközelítések, mint amilyen a döntéselméleti alapú érvelés [4]. A bayesi megközelítés induktív következtetésben, statisztikai következtetésben való tárgyalása a jelen jegyzet Bayesi döntés- és becsléselmélet című fejezetében és az Intelligens adatelemzés című jegyzet több fejezetében található. A szakirodalomból a következő alapművek ajánlhatóak [2, 20, 46] (magyar nyelven lásd [26]), mesterséges intelligencián belüli tárgyalása megtalálható például [3, 47]; történeti áttekintésre és trendek felmérésére pedig a következőek ajánlhatóak [1, 3, 14, 38].
A bayesi értelmezéshez szorosan kapcsolódó mintavételi technikákat röviden a jelen jegyzet Következtetési módszerek című fejezetében foglaltuk össze. Ezeket részletesebben az Intelligen adatelemzés jegyzet több fejezetében is tárgyaljuk. (Általános referenciakiént lásd [8, 18, 20–22, 32, 34, 40]).

2.1.1. Racionális bizonytalanságoktól a valószínűség szubjektív értelmezéséig

Az adat- és tudáselemzésben megjelenő események sokfélesége miatt túlzó egyszerűsítésnek tűnhet a több szinten jelen lévő bizonytalanságot — az adat, a tudás, az elemző,

a modell, az elemzés, az értelmezés bizonytalanságát — ugyanazon formalizmus keretében és ráadásul akár egyetlen keretben kezelni (a rövidség és egyszerűség kedvéért véges, diszkrét eseményrendszerekkel foglalkozunk általában). A keretrendszer univerzális volta miatt feltehető, hogy a bizonytalanság nem az eseményrendszerhez, hanem azon belül az eseményekhez kapcsolódik (eltérően egyéb módszerektől, mint a fuzzy vagy a Dempster-Shaffer megközelítés). Egy ilyen teljes eseményrendszer definiálása, kölcsönösen kizáró és teljes atomi eseményekkel, különösen nehéz olyan szakterületeken, amelyeken eleve több, autonóm, de bizonytalanul is kapcsolódó szint van. Ilyen az orvosbiológa is. Ennek gyakorlására a tárgyhoz kapcsolódó laborgyakorlatok és anyagok adnak lehetőséget.
Egy helyes eseményrendszert feltételezve az események bizonytalanságának értelmezésére több megközelítés is népszerűvé vált a valószínűségszámítás történetének évszázadai során. Ilyen többek között a szerencsejátékokhoz köthető kombinatorikus értelmezés, a fizikalista avagy propensity alapú megközelítés [44], a frekventista, sorozatok határértékeként való értelmezés (az eredeti, von Misses-hez köthető megközelítés áttekintésére és annak számításelméleti aspektusainak kiterjesztésére lásd [54]). A bizonytalanság formalizálására és a valószínűségek értelmezésére unikális lehetőséget nyújt az axiomatikus megközelítés, amely leegyszerűsítve az események bizonytalanságai feletti logikai (preferencia) reláció feltevéséből bizonyítja be a létezését és egyértelmű voltát a valószínűségszámítás Kolmogorov-féle halmazelméleti megalapozásával kompatibilis „pontszámnak", racionalitási axiómák (meta-axiómák) feltevésével. Az így származtatott pontszám valószínűségi mértékk;nt használható, amelynek értelmezése a racionalitási axiómákból következően döntéselméleti alapú, de nevezett szubjektív, bayesi, személyes valószínűségi értelmezésnek is (a teljeskörű tárgyalásért lásd [4]). Az axiomatikus megközelítéshez közelinek mondható a pragmatista vagy eszközhasználati értelmezés, amely a valószínűségi megközelítés modellezésben való felhasználását helyezi az értelmezés középpontjába [7,12,20]. Megközelítésünkben a szubjektív-pragmatista értelmezést követjük, amely ismételt, valamiféle állandóságot mutató megfigyelések esetén a fizikalista vagy frekventista értelmezéssel is összhangba hozható, azok ontológiai elkötelezettsége nélkül.
A bayesi értelmezés bemutatására sorra vesszük a döntéselméleti alapú axiomatikus származtatás főbb lépéseit. A feltevések egy racionalista szubjektum modellezésének szemszögéből is értelmezhető .

8. definíció ([4]). A döntési problémát $\mathcal{E}, \mathcal{C}, \mathcal{A}, \leq$, definálja, ahol:

(i) $\mathcal{E}$ algebrája az eseményeknek, E_j;

(ii) $\mathcal{C}$ a halmaza a lehetséges következményeknek, c_j;

(iii) $\mathcal{A}$ a halmaza a lehetséges cselekedeteknek, amelyek események partícióit feleltetik meg következményeknek;

(iv) $\leq$ egy bináris preferenciareláció $\mathcal{A}$ elemei felett.

Amint a definíció mutatja az eseményrendszer feletti bizonytalanságot csupán egy preferencia reláció testesíti meg. A preferencia reláció értelmezése kulcsfontosságú: definíció szerint aktív beavatkozáshoz kötődik, de szándékolt (és később bebizonyosodó)

jelentése „valószínűbb”. A preferenciarelációból összehasonlíthatósági, tranzitivitás és mennyiségi-folytonossági feltevésekkel az alábbi eredmény bizonyítható.

1. Propozíció. [[4]] Adott $\leq$ bizonytalansági reláció esetén egyértelműen létezik egy valós szám P(E) minden eseményhez, s ezek a számok kompatibilisek $\leq$-val (azaz $E \leq F \Leftrightarrow P(E) \leq P(F)$) és véges, additív valószínűségi mértéket alkotnak.

Egy párhuzamos eredmény korlátos következmények esetén származtatja a preferencia relációval kompatibilis (azaz azt tökélesen modellező) hasznosságok (veszteségek) egyértelmű létét.

2. Propozíció. [[4]] Egy döntési problémában $\mathcal{E}, \mathcal{C}, \mathcal{A}, \leq$ korlátos következmények esetén $c_* < c^*$,

(i) minden c-re a *hasznosság függvény* $u : \mathcal{C} \to \mathcal{R}$ $u(c) = u(c|c_*, c^*)$ létezik és egyértelmű;

(ii) az érték $u(c|c_*, c^*)$ független valamely esemény G feltételezett bekövetkeztétől;

(iii) $0 = u(c_*|c_*, c^*) \leq u(c|c_*, c^*) \leq u(c^*|c_*, c^*) = 1$;

(iv) teljesül az úgynevezett *maximális hasznosság elve* :

$$a_1 \leq_G a_2 \Leftrightarrow \sum_j u(c_{a_1(\mathrm{E}_j)}) P(\mathrm{E}_j\,|G) \leq \sum_j u(c_{a_2(\mathrm{E}_j)}) P(\mathrm{E}_j\,|G) \tag{2.1}$$

Bár a döntéselméleti keret és a származtatáshoz használt „racionalitási” axiómák megkérdőjelezhetők, más axiomatikus alapok is hasonlóan a valószínűségelméletben megszokott mértékeket származtatják, amelyek kompatibilisek, pontosabban egybeesnek az eseményekhez tartozó bizonytalanságok közötti relációkkal, így végeredményként normatívan írják elő a valószínűségi modellezés használatát. A fenti eredmények kiterjeszthetőek komplex események feletti szigma-algebrákra, amely esetben szigma-additív valószínűségi mértékek adódnak eredményként, a valószínűségelmélet Kolmogorov-féle felépítésének megfelelően [45].

2.1.2. Felcserélhetőségtől a bayesi modellátlagolásig

A bizonytalanságok reprezentálásának és értelmezésének axiomatikus megközelítése megmutatta, hogy a racionalitási axiomákat elfogadva az események feletti bizonytalanságok, sőt döntések feletti preferenciák modellezésére is normatívan adódik a valószínűségszámítás és valószínűségi döntéselmélet. Az autonóm, racionális entitás pillanatnyi állapotának a reprezentálását meghaladó induktív esethez azonban tovább kell lépni, ami egy adott stabilitású megfigyelések esetében a valószínűségi keret szükségszerű kiterjesztéséhez vezet. Ennek bemutatásához idézzük a de Finetti-től származó, bináris eseményekre kimondott reprezentációs tételt, amely általánosabb esetekre is kiterjeszthető.

3. Propozíció. [[4]] Ha $x_1, x_2, \ldots$ egy végtelen felcserélhetőségű 0-1 véletlen sorozat P valószínűségi mérték szerint, azaz bármely n-re és permutációra $\pi(1), \ldots, \pi(n)$ az együttes valószínűségi tömegfüggvény $p(x_1, \ldots, x_n) = p(x_{\pi(1)}, \ldots, x_{\pi(n)})$, akkor létezik

egy olyan eloszlásfüggvény $\mathcal{Q}$, amelynek segítségével $p(x_1, \ldots, x_n)$ felírható a következő alakban

$$p(x_1, \ldots, x_n) = \int_0^1 \prod_{i=1}^{n} \theta^{x_i}(1-\theta^{1-x_i})\,\mathrm{d}\mathcal{Q}(\theta),$$

Itt

$$\mathcal{Q}(\theta) = \lim_{n\to\infty} P[y_n/n \leq \theta],$$

illetve $y_n = x_1 + \cdots + x_n$ és $\theta = \lim_{n\to\infty} y_n/n$.

Az úgynevezett reprezentációs tétel szerint a végtelen felcserélhetőség feltevése, nevezetesen, hogy a szubjektív, bayesi valószínűségek a megfigyelés sorrendjétől függetlenek, azt vonja maga után *mintha* a megfigyelések egymástól feltételesen függetlenek lennének egy hipotetikus mintavételi eloszlás paraméterével vett feltétel esetében. Ezen paraméter felett is, mintegy magasabb szinten, megjelenik egy eloszlás, amely a lehetséges paraméterekre mint aszimptotikus határértékekre vonatkozó elvárásokat reprezentálja. Ámbár a végtelen felcserélhetőség az eredmény aszimptotikus volta miatt (a bayesi megközelítésben éppen olyan fontos) véges esetben kritizálható, analóg reprezentációs eredmények szolgáltatják az axiomatikus megközelítés induktív kiterjesztéséből származó Bayes-statisztikai keretet, amelyben az események feletti bayesi, szubjetív valószínűségek egy modellparaméterezés feletti bayesi, szubjetív valószínűségek által indukáltak. (Ennek véges megfigyelések esetén való relevanciájával kapcsolatban észrevehető, hogy bizonyos véges felcserélhetőségű sorozatoknak nincsen keverékreprezentációja, lásd 226.o., [4]).
Ezen eredmények összekapcsolása szerint tehát az események (kimenetelek) feletti bizonytalanság valószínűségekkel reprezentálható, s ez a valószínűségi eloszlás maga is parametrikus eloszlások keverékével reprezentáható (ahol az eloszlások feletti eloszlást az eloszlások paraméterei feletti eloszlás definiálja). Máshogyan fogalmazva: az axiomatikus megközelítések eredményei (például de Finetti, Cox–Jaynes, Bernardo eredményei) arra utalnak, hogy a bizonytalanság kezelésénél a valószínűségszámítás standard, additív elmélete szükségszerűen alkalmazandó, különben a rendszer, ágens, szubjektum egy játékelméleti szituációban veszteséget szenved el - [47]. Pontosabban, ahogyan de Finetti „mintha” tételei mutatják:

I. : A bizonytalanságok feletti racionális (konzisztens) preferencia rendszer valószínűségekkel leírható.

II. : A kimenetelek (akciók) feletti racionális (konzisztens) preferencia rendszer hasznosságokkal (és a maximális hasznosság elvével) leírható.

III. : A felcserélhetőség feltevése maga után vonja a modellátlagolással való leírhatóságot.

Fontos hangsúlyozni, hogy az általunk követett bayesi-pragmatista megközelítésben a valószínűségi keret univerzális alkalmazása mint modellezési eszköz jelenik meg, s az értelmezés semmiben nem érinti a valószínűségszámítás megszokott axiómáit.

2.2. A Bayes-statisztikai keretrendszer általános sémája

Az axiomatikus-pragmatista megközelítésben adódó, megfigyeléseket is kezelő statisztikai keretrendszerben a bizonytalanságok modellezése technikailig két szintre bontható feladatként fogalmazható meg: az események feletti bizonytalanságokat kifejező valószínűségi modell megalkotása, illetve ezen modell feletti bizonytalanságok modellezése egy másik valószínűségi modellel. Fontos hangsúlyozni, hogy csak technikailag elválasztott a két modell és modellezési szint, a kettő egyetlen, egységes valószínűségszámítási keretben jelenik meg (a „valószínűségek valószínűségeivel" kapcsolatos filozófiai vagy pszichológiai megfontolások matematikai oldaról nem jelennek meg). Mielőtt megvizsgálnánk a valószínűségi gráfos modellek felhasználását ezen technikailag kettős szintű keretrendszerben, összefoglaljuk a Bayes-statisztikai keret általános sémáját. A séma gyakorlati aspektusait fogjuk hangsúlyozni, a lényegének bemutatása végett nem tárgyaljuk a döntéselméleti kapcsolatát (azaz itt csak a Bayes-statisztikai keretet vázoljuk, az általános bayesi döntéselméleti keretet a Becslés- és döntéselmélet című fejezetben tárgyaljuk).

Az axiomatikus származtatás jogosságának elfogadásától függetlenül a Bayes-statisztikai keretrendszer koncepcionálisan nagyon egyszerű, s ez a gyakorlati alkalmazásához szükséges számítási erőforrások megjelenése mellett a népszerűségére is magyarázat. Ebben a statisztikai megközelítésben, parametrikus modelleket feltételezve, egy adott információs ellátottságú ξ szituációban a megfigyelések feletti $p(x|\xi)$ bizonytalan elvárásokat úgy állítjuk elő, hogy első lépésként meghatározzuk a releváns, θ paraméterezésű $p(x|\theta)$ modelleket, majd ezen θ paraméterezés felett egy $p(\theta|\xi)$ valószínűség eloszlást (az x_i mennyiségek a megfigyelhető, a θ paraméter a tipikusan nem megfigyelhető kategóriába esnek). A ξ információs kontextus és a valószínűségek feltételeiben való szerepeltetése a valószínűségek szubjektív értelmezését hivatott hangsúlyozni. Gyakran használt jelölés a ξ^+ és ξ^-, amelyek a neminformatív és informatív szituációkat jelölik. A $p(x, \theta|\xi)$ együttes eloszlás megkonstruálása után a valószínűségszámítás szabályai szerint tetszőleges következtetések lehetségesek uniform módon használva a megfigyelhető x_i mennyiségeket és a nem megfigyelhető θ paramétereket. A következtetések általános formában egy $p(\alpha(x, \theta)|\xi)$ közvetlen valószínűségi állítást jelentenek, amely felfogható személyes elvárásnak* A Bayes-statisztikai keret erejét sok tekintetben a valószínűségszámítás ezen uniform, megfigyeléseket és modellparamétereket egyöntetűen kezelő, koherens volta adja.

A jegyzetben a valószínűségszámításban megszokott jelölésrendszert használjuk. Nagybetűs változók jelölik a véletlen változókat és kisbetűs megfelelőik az értékeiket, például $P_X(X = x)$, ahonnan egyértelműség esetén az eloszlás jelölését és a véletlen változót magát is elhagyjuk ($P(x)$). Bináris, propozicionális változók esetében sajnos a konvenció szerint a nagybetű gyakran használt a ponált érték jelölésére is, ezért azon esetekben gyakran a teljes kiírást használjuk ($P(A = igaz)$). Ugyanazt a $p(.)$ jelölést használjuk a valószínűségi tömegfüggvényre és sűrűségfüggvényre, és a megnevezéseiket is megkülönböztetés nélkül használjuk. Ha lehetséges, X jelöli a magyarázó, bemeneti,

*Az angolban szinte kizárólagosan használt „belief" bizonytalanságokkal kapcsolatos neutrális volta a magyarban a meggyőződés és elvárás szavakkal adható talán legjobban vissza, de a hit, hiedelem szavakat is használjuk, esetlegesen érzett mellékjelentéseiktől elvonatkoztatva.

avagy független változót, és Y jelöli a kimeneteli, függő avagy válasz változót. Általában $D_N = \{x^{(1)}, \ldots, x^{(N)}\}$ jelöli N teljes adatot, azaz amikor a változók teljes $\underline{V}$ halmazában minden változó értéke ismert. Ha szükséges, akkor a vektorokat aláhúzás, a mátrixokat dupla aláhúzás jelöli.

2.2.1. A modell specifikálása a bayesi keretben

Egy idealizált bayesi megközelítésben a felhasznált modellek körét a lehető legtágabbra lehetne választani, természetesen az információs kontextusban jelen lévő strukturális kényszerek, például szimmetriák figyelembevételével. Azonban három aspektust mindenképpen érdemes megfontolni:az Ockham–elv potenciális megsértését, a számítási komplexitást és a komplex modell specifikációjának gyakorlati nehézségeit. Az Ockham elv („Ochkham borotvája") szembeállítása a bayesi keretrendszerrel sokat vitatott kérdés. A még mindig parázsló vita a gyakorlatban úgy oldódik meg, hogy a bayesi megközelítésben a komplexebb, általánosabb modellek kevesebb megerősítést kapnak, mint a megfigyelésekhez hasonlóan illeszkedő, de egyszerűbb modellek [28, 33, 40]. A második, számítási komplexitással kapcsolatos ellenvetést a 2000-es évektől megfigyelhető hardverfejlesztési trendek részben megválaszolják, mivel az egyre elérhetőbb párhuzamos architektúrák, általános célú grafikus kártyák (GPGPU-k), vagy akár a „közüzemi szolgáltatásként" igénybe vehető felhőinfrastruktúrák, nagyon jól kihasználhatóak a bayesi következtetésben a Monte Carlo eljárásokon belül. A harmadik, a komplex modellek specifikálásának nehézségéhez kapcsolódó ellenvetés a fejezet fő témája, nevezetesen, hogy hogyan is lehet hatékonyan komplex valószínűségi modelleket formalizálni. A keretek kijelölése és a fogalmak bevezetése érdekében most csak a hierarchikus modellezés koncepcióját foglaljuk össze.
Hierarchikus modellek axiomatikus származtatásához szintén a felcserélhetőségi megfontolások használhatóak fel, de ekkor már a θ paraméterek szintjén, ami analóg módon vezet egy újabb (technikai) „hiper" szint megjelenéséhez modellek együttese (keveréke) és a hozzájuk tartozó ϕ hiperparaméterek által:

$$p(\theta, \phi) = p(\phi)p(\theta|\phi). \tag{2.2}$$

A gyakorlatban elterjedt megközelítés szerint a hierarchikus specifikációban a releváns $\mathcal{M}^i$ modellosztályok specifikációjával, majd az azokon belüli $\mathcal{S}_k^i$ vagy M_k^i model struktúrák specifikációjával, és végül a modellstruktúrákhoz tartozó θ_k^i paraméterek specifikációjával történik. Ennek megfelelően egy adott i modellosztálybeli k struktúra θ_k^i paraméterezéséhez tartozó *a priori* bizonytalan elvárás egy szorzatként fejezhető ki:

$$p(\theta_k^i, M_k^i, \mathcal{M}^i) = p(\mathcal{M}^i)p(M_k^i|\mathcal{M}^i)p(\theta_k^i|M_k^i). \tag{2.3}$$

A modellek eloszlásainak specifikációját a megfigyelhető mennyiségekre vonatkozó $p(x|\theta, \phi)$ (avagy $p(x|\theta_k^i, M_k^i)$ feltételes eloszlás egészíti ki a Bayes-statisztikai megközelítéshez tartozó teljes együttes eloszlássá.

2.2.2. A prediktív következtetés

A Bayes-statisztikai keret felhasználására és a fogalmak bevezetése végett foglaljuk össze a főbb következtetéstípusokat (az induktív következtés részletesen az Intelligens

adatelemzés című jegyzetben található). Az a priori elvárások specifikálása lehetővé teszi az a priori prediktív következtetéseket az x megfigyelhető mennyiségek felett:

$$p(x) = \sum_k p(M_k) \int p(x|\theta_k) p(\theta_k|M_k)\, \mathrm{d}\theta_k. \tag{2.4}$$

Az integrálás és/vagy összegzés a modellek felett, akár modellstruktúrák és paraméterezésük felett, a valószínűségszámításbeli vetítésnek (marginalizációnak) felel meg, és a bayesi kontextusban *bayesi modellátlagolásnak* nevezik [20, 25, 35, 37]. Az a posteriori valószínűségi eloszlások formális bevezetését megelőlegezve, egy D megfigyelési adat esetében a *a posteriori prediktív eloszlás* a D adathalmazon vett feltétellel így írható fel:

$$p(x|D) = \sum_k p(M_k|D) \int p(x|\theta_k) p(\theta_k|D, M_k)\, \mathrm{d}\theta_k. \tag{2.5}$$

Ezek az egyenletek jól illusztrálják a Bayes-statisztikai keret megkülönböztető jegyeit a frekventista valószínűségi értelmezéshez kapcsolódó frekventista statisztikai kerethez képest: a bayesi megközelítés egyetlen fixnek tekintett adathalmaz esetében modellek sokasága felett átlagol (azaz nincs sem modellkiválasztás, sem hipotetikus adathalmazok feletti vizsgálat). A fenti bayesi modellátlagolás normatív származtatása ellenére mint technika is széles körben használt, ilyen például regressziós és klasszifikációs modelleknél a „committee" megközelítésés [5]; Bayes-hálóknál [37]. Mivel a modellátlagolás általában analitikusan nem oldható meg, Monte Carlo módszerek használatosak (általános áttekintését lásd [25]).
Fontos hangsúlyozni, hogy a Bayes-statisztikai megközelítésben a predikció és a prediktív eloszlások jelentik a célt, a modellek csupán eszközkén jelennek meg (bár mint láttuk szükségszerű eszközként). Következésképpen az ideális Bayes-statisztikai eredmény a prediktív eloszlás, amely természetesen nem önmagában jelenik meg (jelentődik), hanem a bayesi döntéselméleti keretben kerül felhasználásra és vezet optimális döntésekhez, például a jelentett mennyiségek révén. Az a priori bizonytalanságok $p(\theta)$ és a $p(x|\theta)$ modell felhasználásával a következtetések másik típusát is bemutatjuk.

2.2.3. A parametrikus következtetés és a Bayes-szabály

Az együttes eloszlás, amelyben a megfigyelt mennyiségeknek és a modellparamétereknek egyforma státuszuk van, a paraméterekre való következtetést is lehetővé teszi. A híres *Bayes szabály* felhasználásával a megfigyelt mennyiségekkel vett feltétel szerinti parametrikus következtetés válik lehetővé:

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{\int p(x|\theta)p(\theta)\, \mathrm{d}\theta} \propto p(x|\theta)p(\theta). \tag{2.6}$$

A 2.6 egyenletben $p(\theta)$ az *a priori eloszlás* vagy prior, $p(x|\theta)$ a *mintavételi eloszlás* amely a *likelihood*-ot és az $L(\theta; x)$ a *likelihood függvény*-t is definiálja. A $p(x)$ az *adat marginális likelihood*-ja, amely csupán egy normalizációs konstanst definiál és $p(\theta|x)$ az

a posteriori eloszlás a paraméterek felett, vagy egyszerűen poszterior. Az a 2.6 egyenlet azt is mutatja, hogy a poszterior egyensúlyi állapotot jelent a prior és likelihood között, és a prior konstans volta miatt a megfigyelések növekvő száma esetén a poszteriort a likelihood fogja dominálni, míg a prior hatása elhanyagolhatóvá válik.
A paraméterek a poszteriori eloszlása már a prediktív poszterior eloszlás kapcsán is megjelent a 2.5 egyenletben mint $p(\theta|D)$, $P(M_k|D)$ és $p(\theta_k|D, M_k)$ a D adat megfigyelése után (lásd [4]). Diszkrét modellek esetében a $p(M_k|D)$ poszterior felírható úgy, hogy

$$p(M_k|D) = \frac{p(D|M_k)p(M_k)}{p(D)} \tag{2.7}$$

ahol a *marginális modell likelihood* avagy M_k evidenciája

$$p(D|M_k) = \int p(D|\theta_k, M_k)p(\theta_k|M_k)\,\mathrm{d}\theta_k \tag{2.8}$$

és a *marginális adat likelihood* pedig

$$p(D) = \sum_k p(D|M_k)p(M_k). \tag{2.9}$$

A bayesi megközelítés elsődlegesen elméleti volta ellenére a bayesi elemzés célja gyakran a modellre vagy legalábbis a modell tulajdonságaira történő következtetés, ami a szubjektív értelmezés szerint a prior elvárások megfigyelések szerinti normatív „frissítését" jelenti.

2.3. Valószínűségi eloszlások függetlenségeinek rendszere

A Bayes-statisztikai keret nyitva hagyja a felhasználandó valószínűségi modellosztály kérdését, így akár a teljes tárgyterületi vagy akár egyetlen változó függésének a modellezése is lehetséges. A továbbiakban a teljes tárgyerületi modellezésnél legelterjedtebb valószínűségi gráfos modelleket, azon belül a Markov-hálókat, s különösen az oksági modellezésben is központi szerepet játszó Bayes-hálókat fogjuk áttekinteni. A feltételes modellek alkalmazhatóságát az Intelligens adatelemzés jegyzet PGM-ek tanulása című fejezetében foglaljuk össze.
A nagyszámú változó feletti együttes eloszlás hatékony reprezentálásánal kulcskérdés az eloszlásban megfigyelhető függetlenségek kihasználása, amihez pedig az explicit reprezentálásuk szükséges. A reprezentációs lehetőségek megértéséhez áttekintjük a függetlenségek rendszerét, s annak szabályszerűségeit.

2.3.1. A függetlenség és feltételes függetlenség fogalmai

A feltételes függetlenség fogalma központi szerepet játszik a valószínűségszámításban, s mint látni fogjuk a relevancia-irrelevencia kérdéskörének tisztázásában a logika, a mesterséges intelligencia, és a gépi tanulás terén is (ezt a valószínűségek szubjektív értelmezése is jelzi). Követve a Dawid [11] által bevezetett jelölést, diszkrét véletlen változók esetében a feltételes függetlenség a következőképpen definiálható.

9. definíció. Legyen $p(\underline{V})$ együttes eloszlás esetén $\underline{X}, \underline{Y}, \underline{Z} \subseteq \underline{V}$ diszjunkt részhalmazok. Jelölje $\underline{X}$ és $\underline{Y}$ $\underline{Z}$ feltétel melletti függetlenségét $I_p(\underline{X}|\underline{Z}|\underline{Y})$, azaz

$$(X \perp\!\!\!\perp Y|Z)_p \text{ } \textit{iff} \text{ } (\forall \underline{x}, \underline{y}, \underline{z} \; p(\underline{x}, \underline{y}|\underline{z}) = p(\underline{x}|\underline{z})p(\underline{y}|\underline{z}) \textit{ ha } p(\underline{z}) > 0). \tag{2.10}$$

Egyéb ekvivalens definíciók még a következők is (feltéve, hogy a szükséges mennyiségek jól definiáltak) [31].

$$\begin{aligned}
(X \perp\!\!\!\perp Y|Z)_p &\Leftrightarrow p(\underline{x}|\underline{y}, \underline{z}) = p(\underline{x}|\underline{z}) \\
(X \perp\!\!\!\perp Y|Z)_p &\Leftrightarrow p(\underline{x}, \underline{y}, \underline{z}) = p(\underline{x}, \underline{z})p(\underline{y}, \underline{z})/p(\underline{z}) \\
(X \perp\!\!\!\perp Y|Z)_p &\Leftrightarrow p(\underline{x}, \underline{y}, \underline{z}) = h(\underline{x}, \underline{z})k(\underline{y}, \underline{z}) \textit{valamely } h,k\textit{-ra} \\
(X \perp\!\!\!\perp Y|Z)_p &\Leftrightarrow p(\underline{x}, \underline{z}|\underline{y}) = p(\underline{x}|\underline{z})p(\underline{z}|\underline{y})
\end{aligned} \tag{2.11}$$

Az $(X \perp\!\!\!\perp Y|Z)_p$ feltételes függetlenségre egy másik jelölés az $I_p(X|Z|Y)$ és az $I_p(X;Y|Z)$. Egyértelműség esetén az alsóindexet és a feltételt elhagyjuk. A függetlenség hiányát, azaz a függést $(X \not\perp\!\!\!\perp Y|Z)_p$ jelöli. Érdemes észrevenni, hogy a feltételes függetlenség $\underline{Z}$ minden releváns értékére megkövetelt. A feltételes függetlenség gyengébb formája a kontextuális függetlenség, amely esetében a feltételes függetlenség csak adott $\underline{c}$ értékek esetén áll fenn egy diszjunkt $\underline{C}$ halmaznál. A *kontextuális feltételes függetlenséget* $\underline{X}$ és $\underline{Y}$ között $\underline{Z}$ feltétel mellett a c kontextusban $I_p(\underline{X}|\underline{Z}, \underline{c}|\underline{Y})$ jelöli, azaz amikor

$$I_p(\underline{X}|\underline{Z}, \underline{c}|\underline{Y}) \text{ } \textit{iff} \text{ } (\forall \underline{x}, \underline{y}, \underline{z} \; p(\underline{y}|\underline{z}, \underline{c}, \underline{x}) = p(\underline{y}|\underline{z}, \underline{c}) \textit{ ha } p(\underline{z}, \underline{c}, \underline{x}) > 0). \tag{2.12}$$

Az X és Y közötti függés erősségének kvantitatív jelzésére nagyon sok asszociációs mérték ismert. Egy igen általános standard mérték a (feltételes) *kölcsönös információ*

$$MI_p(X;Y|Z) = \mathrm{KL}(p(X,Y|Z)|p(X|Z)p(Y|Z)). \tag{2.13}$$

2.3.2. Egyéb valószínűségszámítási alapfogalmak

A függetlenség fogalmán túl a valószínűségszámítás elemi eszközkészletét fogjuk csak használni, amelynek összefoglalásaként ajánlható az MI Almanach [47]. Tételesen a következőkre lesz szükség.

1. Eseménytér. Elemi és összetett esemény, additivitás.
2. Együttes eloszlás. Diszkrét eseménytér esetén eloszlás táblázatmodell.
3. Vetítés/bővítés. Események feletti „ki-átlagolás/szummázás/integrálás”.
4. Véletlen változó. Várható érték, variancia, medián, módusz.
5. Feltételes valószínűség. Kétváltozós és általános eset.
6. Bayes szabály. Kétváltozós és általános eset, prior, posterior fogalma, "$\propto$" jelölés.
7. Láncszabály. Tetszőleges sorrend melletti alkalmazhatóság.
8. Naiv következtetés. Tetszőleges feltételes eloszlás származtatása az együttes eloszlásból, levezetés.
9. Egyenlőtlenségek. Markov, Csebisev, Cauchy-Schwarz, Jensen

2.3.3. A Markov-takaró, Markov-határ és közvetlen függés fogalmai

A feltételes függetlenség lehetővé teszi egy adott változó szempontjából irreleváns változók definiálását is, méghozzá többváltozós módon, illetve elégséges és szükséges szempontokból is.

2.1. *definíció.* Egy változóhalmazt $MB_P(X_i)$-t X_i Markov-takaró-jónak nevezünk $P(X_1, \ldots, X_n)$ eloszlásban, ha $(X_i \perp\!\!\!\perp V \setminus MB(X_i) | MB(X_i))_P$ (egyértelműség esetén P nem jelölt). A minimális Markov-takarót Markov-határnak nevezzük és $MBo(X_i)_P$ jelöli.

Ha a Markov-takaró egyértelműen létezik, akkor bevezethető egy szimmetrikus páronkénti reláció a Markov-határbeliségre [11]: $MBM(X_i, X_j)_P$ fennáll X_i és X_j között P-ben, ha

$$MBM(X_i, X_j)_P \leftrightarrow X_j \in MBo(X_i)_P \tag{2.14}$$

A Markov-határbeliségen belül definiálható egy szigorúbb kategória is, amelyet *közvetlen függésnek* nevezünk, ha minden diszjunkt $Z \subseteq V$ halmazra $(X \not\perp\!\!\!\perp Y | Z)$ fennáll (ebben az esetben a függés két változó között is létezik, amikor $Z =$, ami nem feltétlenül igaz a Markov-határbeli változópároknál).

2.3.4. A grafoid axiómák

Egy adott valószínűségi eloszlás esetén a feltételes függetlenségek eleget tesznek a következő tulajdonságoknak, amelyek a valószínűség szubjektív értelmezése esetén irrelevancia tulajdonságokként is olvashatók [11, 22].

a Szimmetria: Az irrelevancia szimmetrikus.

$$I_p(\underline{X}; \underline{Y} | \underline{Z}) \; iff \; I_p(\underline{Y}; \underline{X} | \underline{Z})$$

b Dekompozíció: Irreleváns információ része is irreleváns.

$$I_p(\underline{X}; \underline{Y} \cup \underline{W} | \underline{Z}) \Rightarrow \; I_p(\underline{X}; \underline{Y} | \underline{Z}) \; and \; I_p(\underline{X}; \underline{W} | \underline{Z})$$

c Gyenge unió: Irreleváns információ irreleváns marad más irreleváns információ megismerése után is.

$$I_p(\underline{X}; \underline{Y} \cup \underline{W} | \underline{Z}) \Rightarrow \; I_p(\underline{X}; \underline{Y} | \underline{Z} \cup \underline{W})$$

d Összevonás: Irreleváns információ irreleváns marad más irreleváns információ elfelejtése után is.

$$I_p(\underline{X}; \underline{Y} | \underline{Z}) \; and \; I_p(\underline{X}; \underline{W} | \underline{Z} \cup \underline{Y}) \Rightarrow \; I_p(\underline{X}; \underline{Y} \cup \underline{W} | \underline{Z})$$

e Metszet : Szimmetrikus irrelevancia együttes irrelevanciát jelent, ha nincs más függés.

$$I_p(\underline{X}; \underline{Y} | \underline{Z} \cup \underline{W}) \; and \; I_p(\underline{X}; \underline{W} | \underline{Z} \cup \underline{Y}) \Rightarrow \; I_p(\underline{X}; \underline{Y} \cup \underline{W} | \underline{Z})$$

Ezek a tulajdonságok más területeken, az adatbázisok elméletében, illetve a gráfelméletben is megjelennek, ami indokolta az alábbi fogalmaj bevezetését:

1. Szemi-grafoid (SG) axiómáknak nevezzük a szimmetria, dekompozíció, gyenge unió, összevonás tulajdonságokat, amelyek minden eloszlásban teljesülnek.

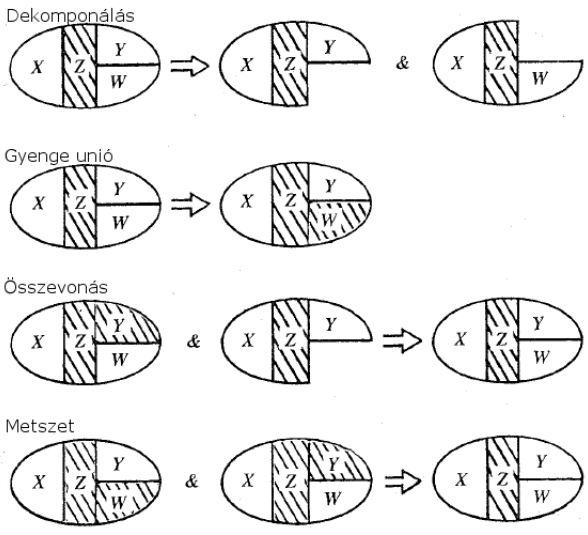
2. Grafoid axiómáknak nevezzük a szemi-grafoid axiómákat és a metszet tulajdonságot, amely csak szigorúan pozitív eloszlásokban áll fenn.

A tulajdonságokat a 2.1 ábra és a 2.2 ábra illusztrálják.

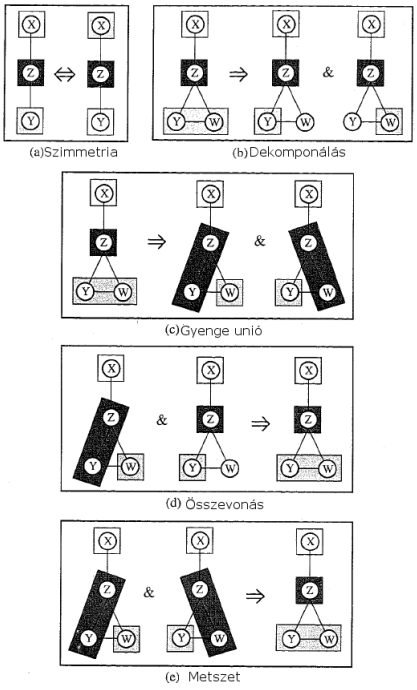
A tulajdonságok bevezetése felvetette, hogy esetleg lehetséges egy helyes és teljes logikai kalkulust létrehozni a függetlenségek feletti következtetésre. Elsőként vezessük be a függetlenségi modell fogalmát.

2.2. *definíció.* Egy $P(X_1, \ldots, X_n)$ eloszlás M_P függetlenségi modellje pontosan a P-ben érvényes $I_P(X, Y|Y)$ függetlenségi állításokat tartalmazza.

Sajnos azonban a Pearl&Paz által1985-ben megfogalmazott teljességi feltevéssel ellentétben a teljesség nem elérhető, mivel vannak ezen kívül is érvényes tulajdonságok, amelyeknek nincsen véges karakterizációjuk [52].



2.1. ábra. A szemi-grafoid axiómák vizualizációja.



2.2. ábra. A grafoid axiómák vizualizációja.

Függetlenségi térképek

Az eloszlások strukturális, sok esetben szinte egyedül lényeges tulajdonságainak, azaz a függetlenségi modelljének a reprezentálásához több megközelítés kínálkozik.

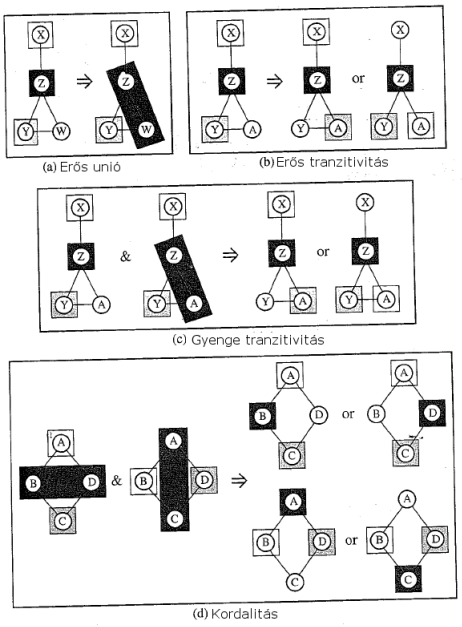
2.3. *definíció.* Egy három argumentumú bináris függvényt $I(X, Y|Y) : V \times V \times V \to 0, 1$ $X, Y, Z \subseteq V$ P eloszlás

1. függetlenségi térképének nevezünk, ha $I \Rightarrow I_p$,

2. függési térképének nevezünk, ha $I \Leftarrow I_p$,

3. perfekt térképének nevezünk, ha $I \Leftrightarrow I_p$.

A grafoid axiómák gráfokban való fennállása miatt a gráfok természetes jelöltek ilyen térkép szerepre, bár korlátaikat látni fogjuk (a 2.3 ábra nem-grafoid axiómákat illusztrál).



2.3. ábra. Nem-grafoid axiómák vizualizációja.

2.4. Valószínűségi gráfos modellek

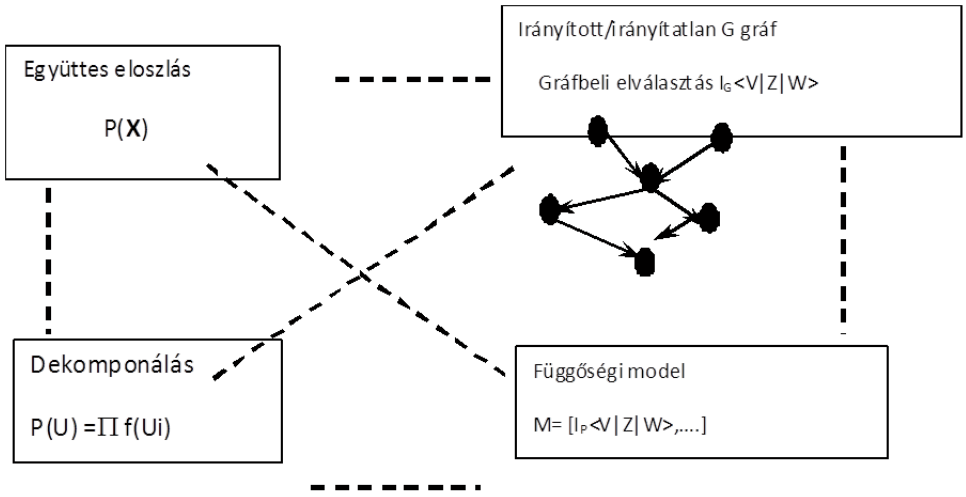
A gráfelmélet és valószínűségszámítás a véletlen gráfok elméletétől a mai hálózatkutatásig sok területen kapcsolódik egymáshoz. Mindkét kutatási irány foglalkozik a valószínűségi eloszlások reprezentálásával, s ez megnehezíti a terminológia választást. Az angol probabilistic graphical models (PGMs) kifejezés jelentése gráfos vagy gráf-alapú valószínűségi-modellek[†], amit a meghonosodni látszó valószínűségi gráfos modellek fordítás remélhetőleg jól kifejez. Mint látni fogjuk, a gráfstruktúrák feletti eloszlások megadásánál a véletlen gráfok elmélete is szerephez jut, de az alapvető kapcsolatot a gráfok felhasználása, a valószínűségi modellek strukturális és parametrikus reprezentálása jelenti. A függetlenségi modell reprezentálásán túl további kérdés az eloszlás kvantitatív reprezentálása, a következtetésben való felhasználása és az oksági modellezésben való felhasználás kérdése. A jelen fejezetben az első két kérdést tárgyaljuk, a következtetés kérdését és az oksági modellek aspektusait a jegyzet két másik fejezete tárgyalja. A jelen fejezetben az oksági kutatásban betöltött szerepe miatt a Bayes-hálós modellosztály kap nagyobb hangsúlyt, de ebben a fejezetben a formális tárgyalás során a valószínűségi (akauzális) értelmezés keretein belül maradunk.

2.4.1. Bayes-hálók kutatásának áttekintése

A Bayes-hálók (BN) a valószínűségi gráfos modellek egy alosztálya, amelyben irányított, körmentes gráfokat (DAG) használunk a sokváltozós eloszlás függetlenségeinek és kvantitatív jellemzőinek a reprezentálására, illetve opcionálisan az eloszlást generáló oksági mechanizmusok reprezentálására is. Egy intuitív értelmezés szerint a csomópontok a véletlen változókat, az élek pedig közvetlen oki ráhatást jelentenek, így definálva

[†]Nem pedig véletlen gráfokon alapuló modellek jelentés.

a modell struktúráját, amely egy irányított körmentes gráf. A struktúrát definiáló DAG lokális valószínűségi modellekkel van kiegészítve, annotálva, nevezetesen minden csomóponthoz tartozik egy lokális modell, amely megadja azon csomópont által reprezentált valószínűségi változó valószínűségi függését a gráfban szülőként jelen lévő valószínűségi változóktól. Ezen lokális modellek paraméterei a modell paraméterei.
A Bayes-háló modellosztályt több tudományterületen is sokoldalúan felhasználják. A teljesség igénye nélkül ide tartozik a tudásmérnökség, a gépi tanulás, az adat- és tudásfúzió, a biomarker kutatások vagy az oksági kutatások. A Bayes-hálók sokoldalúsága abból a tényből következik, hogy három autonóm kutatási szintet kapcsol egybe: az oksági modellt, a valószínűségi modell függetlenségi struktúráját és a kvantitatív eloszlást. Egy másik dimenzió mentén szintén három szerepet tölt be: a tudásreprezentálás, az adatokból történő tanulás és az intelligens döntéstámogatás problémaköreit fedi le. Ezen két dimenzió mentén létrejövő kombinációk kimeríthetetlen tárházat jelentenek, például az a priori ismeretek és megfigyelések, beavatkozások fúziójánál vagy az optimális szekvenciális beavatkozások megtervezésénél. A valószínűségi gráfos modellek és eloszlások akauzális relációinak sokféleségét a 2.4 ábra illusztrálja.



2.4. ábra. A valószínűségi gráfos modellek és eloszlások akauzális reláció.

A gráfos modellek kutatása valószínűségi és oksági modellezésben visszavezethető az 1920-as évekig, Wright útvonal-diagrammokat vizsgáló munkájáig [55]. Az első (orvosi) alkalmazása a Bayes-hálóknak, mint valószínűségi szakértői rendszereknek 1970-ben jelent meg, amely mind tudásmérnöki és gépi tanulási jegyeket is felmutatott [13]. Sokváltozós, nagyobb szakterületet lefedő alkalmazások az 1980-as évek végétől láttak napvilágot. Valószínűségi eloszlások függetlenségeinek szisztematikus vizsgálata 1979-ben publikálták [11], amit J.Pearl mérföldkőnek számító könyve követett a DAG-ok felhasználhatóságáról [22]. Az eloszlások dekomponálásának lehetősége annotált DAG-okkal 1982-ben merült fel először (a gráf alapú dekomponálás részletes tárgyalását lásd [31]).
A Bayes-hálók oksági felhasználása a kezdetektől jelen van [22,28,53], bár eleinte csupán emberi segédeszköznek tekintették és az okság valószínűségi alapú kutatása fő kérdésének a passzív megfigyelésből való tanulás határainak a tisztázását tartották (például

a hatáserősség identifikálhatóságának kérdését [21, 22, 42]). Ezt később egészítette ki a függetlenségi modell mögötti oksági modell kutatása és a kontrafaktuálisok szemantikájának modell alapú definiálása [12, 23]. Hatékony következtetési eljárásokról szóló publikáció Bayes-hálók egy speciális osztályára, a polifákra 1983-ban jelent meg, amelyet 1988-ban követett egy általános esetben is használható, az egzakt következtetésben dominánssá váló megoldás, a klikkek fájában való következtetés [25]. A paraméterek bayesi kezelése rögzített struktúra esetén Dirichlet priorok felhasználásával 1990-ben jelent meg [27], a kapcsolódó „prekvenciális" (predictive sequential,prequential) keret pedig 1993-ból származik [26]. A paraméterek átfogó, strukturális és kauzális aspektusait is figyelembe vevő megoldása 1995-ben született meg [16].

A struktúrák feletti bayesi megközelítés 1991-ig vezethető vissza, amely még feltette a változók oksági sorrendjének az ismeretét [1]. Az általános elméleti és gyakorlati keret 1992-ből származik [6]. Egy teljes bayesi megközelítést strukturális modelltulajdonságok következtetésére 1995-ben közöltek [37], amelyet 2000-ben adaptáltak nagyszámú változóra [11, 16].

Bayes-hálók dekomponált reprezentálása már az 1990-es évektől jelen van a szakirodalomban [19], bár kezdetben a kontextuális függetlenségek vezérelték ezt az irányt. A propozicionális reprezentációtól való elmozdulás a relációs és az általánosabb elsőrendű logika felé egy jelenleg is aktívan kutatott irány [23, 27, 29, 30].

2.4.2. Irányított elválasztás, és egyéb gráfelméleti fogalmak

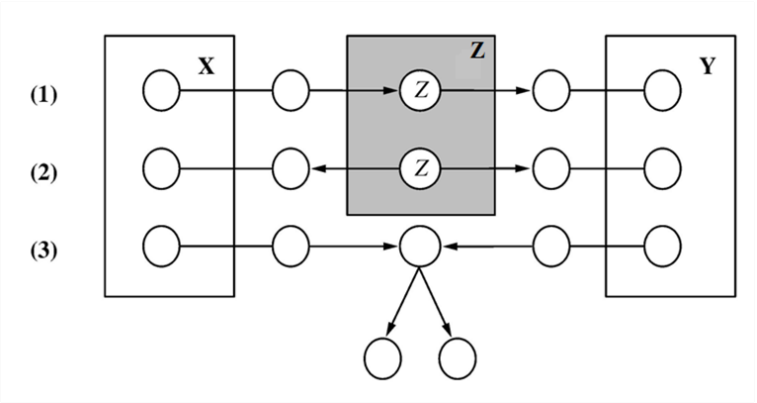
A függetlenségek tulajdonságainak gráfszerűsége és a függetlenségi térképre vonatkozóan megfogalmazott kívánalmak a 2.3 definícióban természetes módon vetik fel a gráfok felhasználását a feltételes függetlenségek reprezentálására. A lefogás/elválasztás reláció lehet az intuitív jelölt erre. Irányított gráfokban ennek a következő definícióját fogjuk használni.

10. definíció. Egy G irányított, körmentes gráfban az $X, Y, Z \subseteq V$ diszjunkt csomópont halmazok esetében jelölje $I_G(X|Z|Y)$, illetve $I_G(X;Y|Z)$, ha X és Y *d-elválasztottak* Z által, azaz ha minden p út X és Y között blokkolt Z által a következőképpen

1,2 a p út tartalmaz egy Z-beli n csomópontot nem összetartó élekkel (azaz így $\rightarrow n \rightarrow$ vagy így $\leftarrow n \rightarrow$),

3 a p út tartalmaz egy nem Z-beli n csomópontot összetartó élekkel (azaz így $\rightarrow n \leftarrow$), amelynek nincs leszármazottja Z-ben.

A 18 definíciót a 2.5 ábra illusztrálja.



2.5. ábra. Az irányított elválasztás (d-szeparáció) vizualizációja. Az X és Y csomóponthalmazok akkor d-szeparáltak, ha köztük minden út blokkolt vagy áthaladó (1) és széttartó (2) módon egy Z-beli csomóponttal, vagy konvergáló módon (3) Z-n kívüli elemmekkel (leszármazottak sem Z-beliek).

Az elválasztáson/lefogáson kívül szükséges gráfelméleti fogalmak még a következőek (lásd MI Almanach [47]):

- Irányítatlan, irányított, részlegesen irányított gráf fogalma
- Gráfreprezentációk, adjacencia mátrix
- Szülő, gyermek, ős, leszármazott, út, hurok, klikk
- Fa, polifa, többszörösön összekötött gráfok
- Topológiai (ősiségi) sorrendezés
- Kordális/háromszögesített gráfok [perfekt gráfok].

Az irányított Markov-feltételek

Egy Bayes-háló struktúrája és a reprezentálni kívánt eloszlás közti kapcsolatot az alábbi négy feltételre alapozhatjuk [7, 22, 23, 31].

11. definíció. A $p(X_1, \ldots, X_n)$ eloszlás faktorizálható a G DAG szerint, ha

$$p(X_1, \ldots, X_n) = \prod_{i=1}^{n} p(X_i | \operatorname{Pa}(X_i)), \tag{2.15}$$

ahol $\operatorname{Pa}(X_i)$ az X_i csomópont szülői halmaza G-ben.

12. definíció. A $p(X_1, \ldots, X_n)$ eloszlásra teljesül a sorrendi Markov-feltétel G szerint, ha

$$\forall\, i = 1, \ldots, n : (X_{\prec(i)} \perp\!\!\!\perp \{\{X_{\prec(1)}, \ldots X_{\prec(i-1)}\} \setminus \operatorname{Pa}(X_{\prec(i)})\} | \operatorname{Pa}(X_{\prec(i)}))_p, \tag{2.16}$$

ahol $\prec$ egy topologikus sorrend G esetén (azaz az élek kompatibilisek G-vel) és $\{X_{\prec(1)}, \ldots X_{\prec(i-1)}\} \setminus \operatorname{Pa}(X_{\prec(i)})$ $X_{\prec(i)}$ összes ősét jelöli kivéve a szüleit.

13. definíció. A $p(X_1, \ldots, X_n)$ eloszlásra teljesül a lokális (szülői) Markov-feltétel G szerint, ha bármely változó független a nem-leszármazottaitól feltéve a szüleit

$$\forall\, i = 1, \ldots, n : (X_i \perp\!\!\!\perp \text{Nondescendants}(X_i) | \,\text{Pa}(X_i))_p, \tag{2.17}$$

ahol Nondescendants(X_i) jelöli X_i nem-leszármazottait G-ben (azaz akikhez nem vezet út G-ben X_i-től).

14. definíció. A $p(X_1, \ldots, X_n)$ eloszlásra teljesül a globális Markov-feltétel G szerint, ha

$$\forall\, X, Y, Z \subseteq V :\ I_G(X; Y|Z)_G \Rightarrow (X \perp\!\!\!\perp Y|Z)_p. \tag{2.18}$$

Ezen feltételek segítségével megfogalmazható egy alapvető kapcsolat eloszlások és DAG reprezentációjuk között [31].

2.4.1. Tétel. [[31]] Egy $p(V)$ eloszlás és G DAG esetén a 11, 12, 13 és 14 feltételek ekvivalensek:

(F) p Markovi G-hez avagy p faktorizálódik G szerint,

(O) p eleget tesz a sorrendi Markov-feltételnek G szerint,

(L) p eleget tesz a lokális Markov-feltételnek G szerint,

(G) p eleget tesz a globális Markov-feltételnek G szerint.

A feltételek ekvivalenciája miatt ezekre a feltételekre együttesen is mint irányított Markov-feltételekre hivatkozhatunk (p, G) pár viszonylatában.

2.4.3. Bayes-háló definíciók

A Markov-feltételek felhasználásával az alábbi meghatározás adható.

15. definíció. A G irányított körmentes gráf a $P(V)$ eloszlás Bayes-hálója, ha minden változót a gráf egy csomópontja reprezentál, a gráfra teljesül valamelyik (és így az összes) Markov-feltétel, és a gráf minimális (azaz bármely él elhagyásával a Markov-feltétel már nem teljesül).

Míg ez a definíció egyértelműen a valószínűségi függetlenségek rendszerének reprezentációjaként tekint a Bayes-hálóra, addig a mérnöki gyakorlatban közkedvelt az alábbi, praktikus meghatározás.

16. definíció. A V valószínűségi változók Bayes-hálója a (G, θ) páros,ha G egy irányított körmentes gráf, amelyben a csomópontok jelképezik V elemeit, θ pedig a csomópontokhoz tartozó $P(X_i|Pa(X_i))$ feltételes eloszlásokat leíró numerikus paraméterek összessége.

A Markov-feltétel teljesülése biztosítja, hogy minden gráfból kiolvasott függetlenség teljesüljön az eloszlásban, azonban a másik irányhoz, ahhoz tehát, hogy minden függetlenség kiolvasható is legyen a gráfból, annak stabilnak is kell lennie.

17. definíció. Egy P(U) eloszlás stabil, ha létezik olyan G DAG, hogy P(U)-ban pontosan a G-ből d-szeparációval kiolvasható függések és függetlenségek teljesülnek benne (azaz G perfekt térkép).

A DAG-reprezentáció korlátját alapvetően az jelenti, hogy numerikusan a struktúra szerint nem szükségszerű függetlenségek is lekódolhatóak. A triviális redundanciákon túl ezek rejtett formákban is megjelenhetnek, például nem tranzitív függések képében vagy alacsonyabbrendű függetlenségek képében (például egy Markov-láncban megfelelő paraméterezés mellett előfordulhat, hogy a függések nem tranzitívak).
A d-szeparáció szükséges és elégséges voltát a következő tétel mutatja, amely szerint egy adott G DAG-gal kompatibilis összes eloszlásban érvényes függetlenségeknek a G-beli d-szeparáció egzakt reprezentációja [22].

2.4.2. Tétel. [[22]]

$$\forall\, X, Y, Z \subseteq V:\ (X \perp\!\!\!\perp Y|Z)_G \Leftrightarrow ((X \perp\!\!\!\perp Y|Z)_p \text{ in all } p \text{ Markov relative to } G.$$

A tétel élesítése, hogy általános bayesi megközelítésben azon eloszlások mértéke 0, amelyeknek G nem perfekt térképe [21].
Azonban bizonyos típusú függetlenségekhez, például a négy csomópontos gyémánt struktúrához, a Bayes-hálós megközelítés nem alkalmas, s ez a reprezentációs korlát felveti más gráftípusok használatát is.

2.4.4. Markov-hálók

Markov-hálók esetében egy G=(P,E) irányítatlan gráf alapján definiálunk egy függetlenségi térképet a G-beli (irányítatlan) elválasztásra/lefogásra alapozva (P a (csomó)pontok halmaza és E a P-be tartozó pontpárok, élek halmaza).

18. definíció. Egy G irányítatlan gráfban az $X, Y, Z \subseteq V$ diszjunkt csomópont halmazok esetében jelölje $I_G(X|Z|Y)$, illetve $I_G(X;Y|Z)$, ha X és Y *elválasztottak* Z által, ha minden p út X és Y között tartalmaz egy Z-beli elemet.

Az irányított gráfokkal analóg módon irányítatlan gráfoknál is megfogalmazhatóak Markov-feltételek.

2.4.5. Markov-feltételek irányítatlan gráfokban

F Klikk faktorizálás: Ha $P(X_{1:n})$ G klikkjein definiált szorzatként felírható.

P Páronkénti markoviság: Bármely két nem szomszédos változó független egymástól az összes többi változóval vett feltétellel.

L Lokális markoviság: Egy változó független minden más változótól a szomszédaival vett feltétellel.

G Globálil markoviság: Bármely két változóhalmaz független egymástól egy őket elválasztó halmazzal vett feltétellel.

Az irányított gráfoktól eltérően ezek a feltételek általános esetben nem ekvivalensek, bár $G \Rightarrow L \Rightarrow P$. Kordális gráfokban szorosabb a kapcsolatuk, és pozitív eloszlásokban viszont ekvivalensek.

2.1. tétel (Hammersley-Clifford, 1971, unpublished)**.** Pozitív eloszlásokban az irányítatlan Markov feltételek ekvivalensek: $F \Leftrightarrow P \Leftrightarrow L \Leftrightarrow G$.

2.2. tétel. Ha egy G kordális gráf lokálisan (vagy globálisan) markovi a p eloszláshoz, akkor az F tulajdonság is teljesül.

Ennek következményeként is a Markov-háló definíciója így adható meg.

2.4. *definíció.* G irányítatlan gráf a p eloszlás Markov-hálója (azaz G függetlenségi térképe p függetlenségi modelljének), ha G globálisan[‡] markovi ($I_G(X|Z|Y) \Rightarrow I_P(X|Z|Y)$) és G minimális (azaz egy él törlésével ezt a tulajdonságát elvesztené).

Pozitív eloszlásokban a Markov-hálót/Markov-(véletlen) mezőt Gibbs-(véletlen) mezőnek is nevezik, mivel a Hammersley–Clifford elmélet szerint a következő faktorizálható lehetséges.

2.5. *definíció.* Egy eloszlás P_Φ Gibbs mérték, ha faktorokkal (potenciálokkal) definiálható $\Phi = \{\Phi_1(V_1), \ldots, \Phi_K(V_K)\}$:

$$P_\Phi(V) = \frac{1}{Z} \prod_{i=1}^{K} \Phi_i(V_i), \tag{2.19}$$

ahol $V_1, \ldots, V_K \subset V$ és Z egy normalizációs konstans. Gibbs-(véletlen) mezőkhöz tartozó G esetében $V_1, \ldots, V_K$ pontosan G klikkjei.

Mivel a globális markoviságból következik a lokális markoviság (p, G) esetében, X szomszédos csomópontjai G gráfban $bd(X, G)$ mindig X Markov-takarója ($MB_p(X, bd(X, G))$). Továbbá pozitív eloszlásokban ez egy egyértelmű Markov-takaró.

2.3. tétel (Pearl,Paz, 1985)**.** Ha G egy Markov-hálója egy p pozitív eloszlásnak, akkor X szomszédos csomópontjai G gráfban egy egyértelmű Markov-határt formálnak.

A tulajdonságok a tudásmérnökség és gépi tanulás számára is sokrétűen felhasználhatóak, például:

F Klikk faktorizálás: következtetésben.

P Páronkénti markoviság: élenkénti konstrukció esetében.

L Lokális markoviság: Markov-határ alapú konstrukcióban.

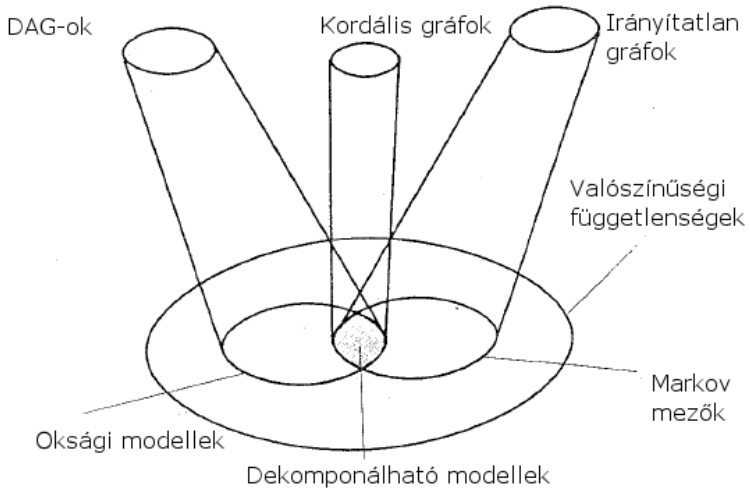
G Globális markoviság: globális relevencia viszonyok gyors kikövetkeztetésében (elválasztásra hatékony algoritmusok léteznek).

[‡]Nem ekvivalens definíciókban a Lokális feltétel is használt.

2.4.6. Bayes-hálók és Markov-hálók reprezentációs képessége

A Markov-hálók, a Bayes-hálókhoz hasonlóan helyes, de nem teljes reprezentációi a függetlenségeknek. Reprezentációs képességeik (avagy hiányosságaik) azonban bizonyos mértékig komplementerek. Például a gyémántstruktúrájú Markov-hálókkal egzakt módon reprezentált eloszlások Bayes-hálókkal egzakt módon nem reprezentálhatóak, viszont a Bayes-hálókban egzakt módon reprezentálható v-struktúra egzakt módon nem reprezentálható Markov-hálókkal. A két modellosztály reprezentációs képességeit a 2.6 ábra illusztrálja.

2.6. *definíció* (Intranzitív hármas/v-struktúra). X, Y, Z véletlen változók egy Intranzitív hármast/v-struktúrát alkotnak p eloszlásbn, ha fennáll $D_p(X;Y)$, $D_p(Y;Z)$ és $I_p(X;Z)$.



2.6. ábra. A különböző típusú gráfok reprezentációs képességei.

2.5. Egyszerű Bayes-hálók

A valószínűségi gráfos modellek strukturális képességeinek áttekintése után a modellek kvantitatív specifikálásának kérdéseit vizsgáljuk. Ennek bevezetéseként három egyszerű, de széleskörben használt modellosztály paraméterezését és felhasználását mutatjuk be.

2.5.1. Naiv Bayes-hálók

Tételezzük fel, hogy a változók halmaza tartalmaz egy hipotézisváltozót (Y) és megfigyeléseket ($X_1, \ldots, X_n$). (Elterjedt más elnevezések ugyanezen fogalmakra: ok, modell, diagnózis, illetve okozat, bizonyíték, megfigyelés, tünet, szimptóma.) A megfigyelések csoportját tekinthetjük különböző típusú megfigyeléseknek (például tüneteknek egy orvosi problémában) vagy azonos típusú, de eltérő idejű megfigyelések szekvenciáinak. A naiv Bayes-hálók definiáló tulajdonsága, hogy a $P(Y, X_1, \ldots, X_n)$ eloszlásban X_i-k feltételesen függetlenek Y feltétellel, azaz $I_P(X_j; X'|Y) = P(X_j|Y)$ bármely diszjunkt X'

részhalmazra. A modell parametrikus megadását Y $p(Y)$ eloszlásának megadása és a $P(X_1|Y),\ldots,P(X_n|Y)$ feltételes eloszlások megadása jelenti. Ekkor

$$P(Y|X_{i_1},\ldots,X_{i_m}) \propto P(X_{i_1}|Y)\ldots P(X_{i_m}|Y)P(Y). \tag{2.20}$$

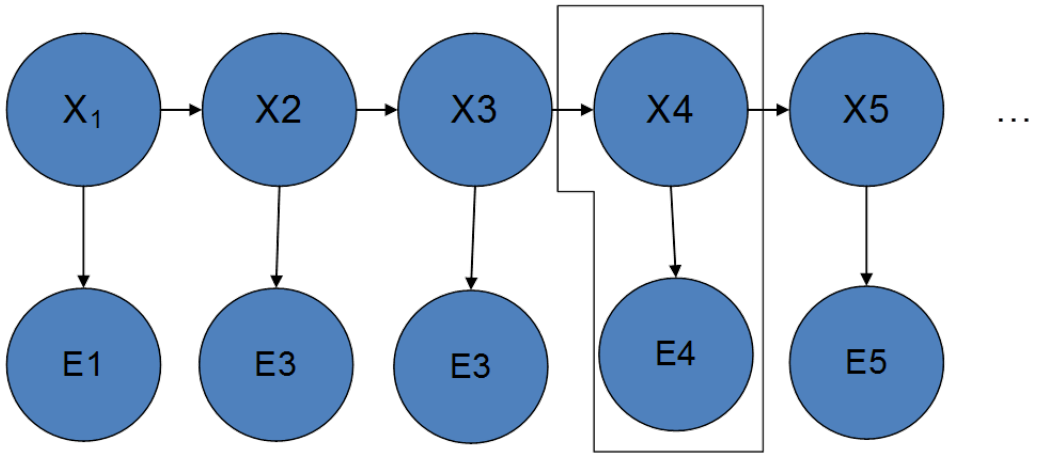
Ha Y bináris, akkor az esély így írható fel:

$$\frac{P(Y=1|X_{i_1},\ldots,X_{i_m})}{P(Y=0|X_{i_1},\ldots,X_{i_m})} = \frac{P(Y=1)}{P(Y=0)}\prod_i \frac{P(X_{i_1}|Y=1)}{P(X_{i_1}|Y=0)}. \tag{2.21}$$

Az eredmény jól tükrözi a valószínűségi gráfos modellek, nevezetesen itt a Bayes-hálók azon képességét, hogy az általános diszkrét, véges esetben exponenciális számosságú paramétert és következtetést lineáris számosságú paraméterre és lineáris számítási idejű következtetésre lehetett redukálni a függetlenségek kihasználásával.

2.5.2. Markov-láncok és rejtett Markov modellek

Elsőrendű Markov-láncok esetében feltételezzük, hogy az $X_1,\ldots,X_n$ valószínűségi változók körében minden i-re teljesül, hogy X_i feltételesen független az $X_1,\ldots,X_{i-2}$ változóktól az X_{i-1} ismeretében. Homogén lánc esetében továbbá az úgynevezett átmeneti valószínűségeket egy indexfüggetlen $P(X_i|X_{i-1})$ feltételes eloszlás definiálja. Ez a modell a rejtett Markov-modelleknél kiegészül a megfigyelhető evidenciák $E_1,\ldots,E_n$ valószínűségi változó halmazzal, amit – egy statikus érzékelő modellt feltételezve – egy szintén indexfüggetlen $P(E_i|X_i)$ köt a közvetlenül már nem megfigyelhetőnek vélelmezett „rejtett állapot” X_i változókhoz. A Markov-láncot és rejtett Markov modelleket a 2.7 ábra illusztrálja.



2.7. ábra. Markov-lánc és rejtett Markov modell illusztrálása.

A rejtett Markov-modellek esetében szintén hatékony, a változók számában lineáris és a változók értékkészletének méretében négyzetes futási idejű eljárások léteznek a következő tipikus feladatokra:

1. Szűrés (filtering) vagy ellenőrző megfigyelés (monitoring): ez a bizonyossági állapot (belief state) kiszámításának a feladata, ami a jelenlegi állapot feletti a

posteriori eloszlás, az adott időpontig vett összes bizonyíték ismeretében; vagyis szeretnénk kiszámítani a $P(X_t|e_{1:t})$ mennyiséget, feltéve, hogy a bizonyítékok folyamatos sorozatban érkeznek a $t = 1$ időponttól kezdve.

2. Előrejelzés (prediction): ez egy jövőbeli állapot feletti a posteriori eloszlás kiszámításának a feladata az adott időpontig vett összes bizonyíték ismeretében; azaz, szeretnénk kiszámítani a $P(X_{t+k}|e_{1:t})$ mennyiséget valamely $0 < k$ esetén.

3. Simítás (smoothing) vagy visszatekintés (hindsight): ez egy múltbeli állapot feletti a posteriori eloszlás kiszámításának a feladata a jelen időpontig vett összes bizonyíték ismeretében; azaz, szeretnénk kiszámítani a $P(X_k|e_{1:t})$ mennyiséget valamely $0 < k < t$ esetén.

4. Legvalószínűbb magyarázat (most likely explanation): a megfigyelések egy sorozatának ismeretében szeretnénk megtalálni azt az állapotsorozatot, amely a legvalószínűbben generálta az adott megfigyeléseket; vagyis szeretnénk kiszámítani az $argmax_{x_{1:t}} P(x_{1:t}|e_{1:t})$ értékét.

5. Likelihood paraméterezés: megfigyelések egy vagy több sorozatának ismeretében (x) szeretnénk megtalálni azt a paraméterezést (θ), amely az adott megfigyeléseket legnagyobb valószínűséggel generálta; azaz szeretnénk kiszámítani az $argmax_\theta P(x|\theta)$ értékét.

2.6. Parametrizáció, priorok definiálása és tudásmérnöki kérdések

A teljes bayesi megközelítésben a modelleket nem csupán egy pontparametrizációval kell ellátni, hanem a priori kényszereknek megfelelő struktúrák és paramétertér feletti eloszlások megadásával is. Ezek tárgyalása viszont az oksági aspektusok figyelembevételét is igényli, így tárgyalásuk a tudásmérnöki és kiterjesztett PGM reprezentációkhoz hasonlóan az Oksági modellek című fejezetben kapott helyet.

Irodalomjegyzék

[1] C. Andrieu, A. Doucet, and C. P. Robert. Computational advances for and from Bayesian analysis. *Statistical Science*, 19(1):118–127, 2004.

[2] J. O. Berger. *Statistical Decision Theory and Bayesian Analysis.* Springer-Verlag, 1980.

[3] J. O. Berger. Bayesian analysis: A look at today and thoughts of tomorrow. *Journal of the American Statistical Association*, 95(452):1269–1276, 2000.

[4] J. M. Bernardo. *Bayesian Theory.* Wiley & Sons, Chichester, 1995.

[5] C. M. Bishop. *Neural Networks for Pattern Recognition.* Clarendon Press, Oxford, 1995.

[6] W. L. Buntine. Theory refinement of Bayesian networks. In *Proc. of the 7th Conf. on Uncertainty in Artificial Intelligence (UAI-1991)*, pages 52–60. Morgan Kaufmann, 1991.

[7] P. Cheeseman. In defense of probability. In *Proceedings of the Ninth International Joint Conference on Artificial Intelligence (IJCAI-85)*, pages 1002–1009. Morgan Kaufmann, 1985.

[8] M. Chen, Q. Shao, and J. G. Ibrahim. *Monte Carlo Methods in Bayesian Computation.* Springer-Verlag, New York, 2000.

[9] G. F. Cooper and E. Herskovits. A Bayesian method for the induction of probabilistic networks from data. *Machine Learning*, 9:309–347, 1992.

[10] R. G. Cowell, A. P. Dawid, S. L. Lauritzen, and D. J. Spiegelhalter. *Probabilistic networks and expert systems.* Springer-Verlag, New York, 1999.

[11] A. P. Dawid. Conditional independence in statistitical theory. *J. of the Royal Statistical Soc. Ser.B*, 41:1–31, 1979.

[12] A. P. Dawid. Probability, causality and the empirical world: A bayes-de finetti-popper-borel synthesis. *Statistical Science*, 19(1):44–57, 2004.

[13] F. T. de Dombal, D. J. Leaper, J. C. Horrocks, and J. R. Staniland. Human and computer-aided diagnosis of abdominal pain. *British Medical Journal*, 1:376–380, 1974.

[14] B. Efron. Bayes' theorem in the 21st century. *Proc. Natl. Acad. Sci.*, 340(7):1177–78, 2013.

[15] N. Friedman and D. Koller. Being Bayesian about network structure. In *Proc. of the 16th Conf. on Uncertainty in Artificial Intelligence(UAI-2000)*, pages 201–211. Morgan Kaufmann, 2000.

[16] N. Friedman and D. Koller. Being Bayesian about network structure. *Machine Learning*, 50:95–125, 2003.

[17] D. Galles and J. Pearl. Axioms of causal relevance. *Artificial Intelligence*, 97(1-2):9–43, 1997.

[18] D. Gamerman. *Markov Chain Monte Carlo*. Chapman & Hall, London, 1997.

[19] D. Geiger and D. Heckerman. Knowledge representation and inference in similarity networks and Bayesian multinets. *Artificial Intelligence*, 82:45–74, 1996.

[20] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian Data Analysis*. Chapman & Hall, London, 1995.

[21] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter. *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London, 1996.

[22] L. Győrfi. *T omegkiszolgálás informatikai rendszerekben*. Műegyetemi Kiadó, 1996. (in Hungarian).

[23] J. Y. Halpern. An analysis of first-order logics of probability. *Artificial Intelligence*, 46:311–350, 1990.

[24] D. Heckerman, D. Geiger, and D. Chickering. Learning Bayesian networks: The combination of knowledge and statistical data. *Machine Learning*, 20:197–243, 1995.

[25] J. A Hoeting, D. Madigan, A. E. Raftery, and C. T. Volinsky. Bayesian model averaging: A tutorial. *Statistical Science*, 14(4):382–417, 1999.

[26] L. Hunyadi. Bayes gondolkodás a statisztikában. *Statisztikai szemle*, 89(10-11):1150–1171, 2011.

[27] Manfred Jaeger. Relational Bayesian networks. *Proc. of the 13th Conference on Uncertainty in Artificial Intelligence (UAI-1997)*, pages 266–273, 1997.

[28] W. H. Jefferys and J. O. Berger. Sharpening ockham's razor on a Bayesian strop, 1991.

[29] D. Koller and A. Pfeffer. Object-oriented Bayesian networks. In Dan Geiger and Prakash P. Shenoy, editors, *Proc. of the 13th Conf. on Uncertainty in Artificial Intelligence (UAI-1997)*, pages 302–313. Morgan Kaufmann, 1997.

[30] D. Koller and A. Pfeffer. Probabilistic frame-based systems. In *Proc. of the 15th National Conference on Artificial Intelligence (AAAI), Madison, Wisconsin*, pages 580–587, 1998.

[31] S. L. Lauritzen. *Graphical Models.* Oxford, UK, Clarendon, 1996.

[32] J. S. Liu. *Monte Carlo Strategies in Scientific Computing.* Springer-Verlag, 2004.

[33] D. J. C. MacKay. Probable networks and plausible predictions - a review of practical Bayesian methods for supervised neural networks. *Neural Computation*, pages 469–505, 1996.

[34] D. J. C. Mackay. *Learning in graphical models*, chapter Introduction to Monte Carlo Methods. MIT Press, Cambridge, MA, 1999.

[35] D. Madigan and R. Almond. Test selection strategies for belief networks. StatSci Research Report 20., 1993.

[36] D. Madigan, S. A. Andersson, M. Perlman, and C. T. Volinsky. Bayesian model averaging and model selection for Markov equivalence classes of acyclic digraphs. *Comm.Statist. Theory Methods*, 25:2493–2520, 1996.

[37] D. Madigan and J.York. Bayesian graphical models for discrete data. *Internat. Statist. Rev.*, 63:215–232, 1995.

[38] D. Malakoff. Bayes offers a 'new' way to make sense of numbers. *Science*, 286:1460–1464, 1999.

[39] C. Meek. Causal inference and causal explanation with background knowledge. In *Proc. of the 11th Conf. on Uncertainty in Artificial Intelligence (UAI-1995)*, pages 403–410. Morgan Kaufmann, 1995.

[40] R. M. Neal. *Bayesian Learning for Neural Networks.* Springer, Berlin, 1996.

[41] J. Pearl. *Probabilistic Reasoning in Intelligent Systems.* Morgan Kaufmann, San Francisco, CA, 1988.

[42] J. Pearl. Causal diagrams for empirical research. *Biometrika*, 82(4):669–710, 1995.

[43] J. Pearl. *Causality: Models, Reasoning, and Inference.* Cambridge University Press, 2000.

[44] K. R. Popper. *The Logic of Scientific Discovery.* Hutchinson, London, 1959.

[45] A. Rényi. *Probability Theory.* Akadémiai Kiadó, Budapest, 1970.

[46] C. P. Robert. *The Bayesian Choice.* Springer-Verlag, New York, 2001.

[47] S. Russel and P. Norvig. *Artificial Intelligence.* Prentice Hall, 2001.

[48] D. J. Spiegelhalter. Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society, Series B*, 50(2):157–224, 1988.

[49] D. J. Spiegelhalter, A. Dawid, S. Lauritzen, and R. Cowell. Bayesian analysis in expert systems. *Statistical Science*, 8(3):219–283, 1993.

[50] D. J. Spiegelhalter and S. L. Lauritzen. Sequential updating of conditional probabilities on directed acyclic graphical structures. *Networks*, 20(.):579–605, 1990.

[51] P. Spirtes, C. Glymour, and R. Scheines. *Causation, Prediction, and Search*. MIT Press, 2001.

[52] M. Studeny. Semigraphoids and structures of probabilistic conditional independence. *Annals of Mathematics and Artificial Intelligence*, 21(1):71–98, 1997.

[53] T. Verma and J. Pearl. *Causal Networks: Semantics and Expressiveness*, volume 4, pages 69–76. Elsevier, 1988.

[54] G. Vita'nyi and K. Li. *An introduction to Kolmogorov complexity and it applicationes*. Springer-Verlag, New York, 1990.

[55] S. Wright. Correlation and causation. *J. of Agricultural Research*, 20:557–585, 1921.

3. fejezet

Oksági modellek: reprezentációk és következtetések

3.1. Bevezető

A Valószínűségi gráfos modellek című fejezetben ismertett megközelítés a bizonytalanság kezelésére, modellezésére mind tudománytörténeti, mind gyakorlati szempontból sikeresnek mondható, aminek eredményeképpen a bayesi statisztika, a bayesi döntéselmélet, illetve a valószínűségi gráfos modellek, ezen belül a Bayes-hálók és a Markov-hálók megjelentek és széles körben elterjedtek szinte minden tudományterületen, de ipari és kereskedelmi alkalmazásokban is. Az oksági kutatás sok tekintetben eltérő helyzetben van, mivel az okozatiság

1. inkább a determinisztikus és nem bizonytalan világképhez tartozik,
2. aszimmetrikus, szemben az információs, asszociációs bizonytalansággal,
3. aktív cselekvések, beavatkozások következményeihez kapcsolódik, és nem passzív megfigyelésekhez,
4. mechanizmusokhoz kapcsolódik, amelyek autonómok, modulárisok az őket terhelő zajok és a beavatkozások viszonylatában,
5. idői-aspektussal is rendelkezik.

A bizonytalanság modellezésében az asszociációs relációk és az oksági relációk megkülönböztetésére több szempontrendszert is megfogalmaztak, ilyen például az orvosbiológiai kutatásokból származó következő lista, mely az oksági relációkkal szemben támasztott követelményeket sorolja fel [33]:

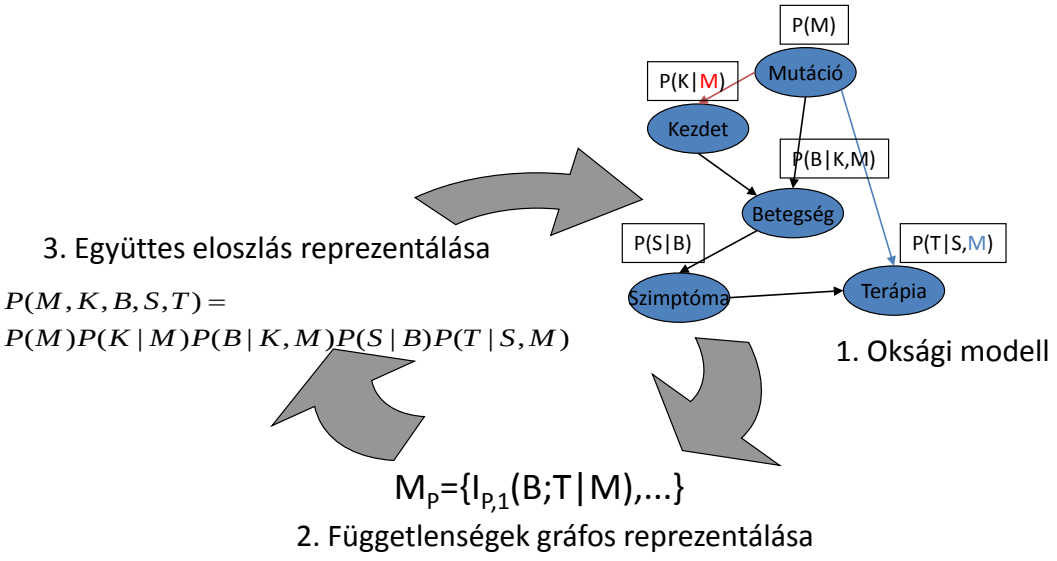
1. Erő. Erős statisztikai asszociáció.
2. Konzisztencia, specifikusság, koherencia. Például az ok megszüntetésével a hatás is szűnjön meg (szükségesség), és az ok bekövetkeztével a hatás is erősödjön (elégségesség).
3. Gradiens. Legyen a következmény arányos a hatással (dózis–hatás elv).

4. Temporalitás. X időben előzze meg Y-t.

5. Plauzibilitás és analógia. Létezzen magyarázat, és ne legyenek alternatív, zavaró tényezőre is építő alternatív magyarázatok.

6. Kísérleti adatok léte.

Jelen fejezetben csak az oksági modellek szakértői tudáson alapuló létrehozását és felhasználását vizsgáljuk, a tanulásukat az Intelligens adatelemzés című jegyzet egyik fejezete tartalmazza.
Kiindulópontként elfogadjuk a Bayes-statisztikai keretet. Célunk egy olyan modell létrehozása, amely összegzi az eddigi ismereteket és megfigyeléseket, s amely lehetővé teszi a beavatkozások hatásainak automatizált kikövetkeztetését is. Elsőként azt vizsgáljuk meg, hogy a Valószínűségi gráfos modellek című fejezetben ismertetett Bayes-hálók mennyiben felelnek meg az oksági modellekkel kapcsolatos intuíciónknak, nevezetesen annak az oksági szemantikának, hogy az élek közvetlen oksági ráhatást reprezentálnak. Az irányított körmentes gráfok, a DAG-ok, remélt hármas felhasználását a 3.1 ábra illusztrálja.



3.1. ábra. Bayes-hálók reprezentációjának három aspektusa.

Az experimentális oldalról közelítve bevezethető az ideális beavatkozáshoz tartozó (empirikus) eloszlás fogalma.

19. definíció. Jelölje $do(X = x)$ az X változó beállítását x értékre, és jelölje $\hat{p}(Y|do(x))$ az ehhez a beavatkozáshoz tartozó eloszlást (elméleti kontextusban az empirikus voltát jelölő $\hat{p}$ nem jelölt).

Természetesen a megfigyelési és beavatkozási eloszlások már a legegyszerűbb esetben is eltérhetnek. Fontoljuk meg a két X, Y változó alkotta rendszert, amelyeket egy $X \to Y$ oksági reláció köt össze indulálva a $p(X, Y)$ eloszlást. Az X esetében nincs különbség (passzív) megfigyelésés és (aktív) beavatkozás között, de Y esetében már igen: $p(Y|do(x)) = p(Y|x)$, de $p(X|do(y)) = p(X)$ és nem egyenlő $p(X|y)$-nal. A megfigyelési ekvivalencia analógiájára a *kauzális irrelevencia* fogalma is bevezethető [12, 23].

A statisztikai és oksági kapcsolatak kutatásának legfőbb fogyasztója epidemiológia. E területen a következő mérõszámok használtak az oksági relációk kvantitatív jellemzésére, amelyek a *do* szemantikát felhasználva a következőképpen írhatóak (egy bináris X (pl. kitettség) és Y (pl. betegség) között): rizikóbeli különbség vagy oksági hatás (δ), tulajdonítható/okozott rizikó (θ) és az esélyhányados(Ψ).

$$\delta = p(y|do(x)) - p(y|do(\neg x) \tag{3.1}$$

$$\theta = \frac{p(y|do(x)) - p(y|do(\neg x)}{p(y|do(x))} \tag{3.2}$$

$$\Psi = \frac{p(y|do(x))/p(\neg y|do(x))}{p(y|do(\neg x))/p(\neg y|do(\neg x))} \tag{3.3}$$

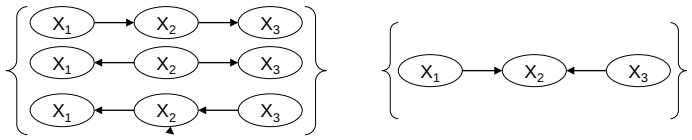
Természetesen ezeknek a mennyiségeknek a standard epidemiológiai definíciója nem a beavatkozásos *do* szemantikát használja, hanem az „adjustált" megfigyelési valószínégéken alapulókat [24, 32, 33]. Az „adjustálás" (vagy „kontrollolálás") a zavaró tényezők $\underline{Z}$ eliminálására szolgál úgy, hogy X hatását Y-ra a potenciális zavaró tényezők azonos értékei mellett vizsgáljuk (azaz feltételbe emeljük és „fixen tartjuk" őket).

$$p(y|do(x)) = \sum_z p(y|x, z)p(z) \tag{3.4}$$

3.2. Bayes-hálók ekvivalencia-osztályai

Az eloszlás stabilitásának és szigorú pozitvitásának feltevése sem zárja ki, hogy az eloszlás függetlenségi modelljének több DAG is perfekt térképe legyen. Erre példa a következő.

3.2.1. Példa. [] Tekintsünk egy Markov-láncot $\mathcal{X} = \{X_1, \ldots, X_n\}$, amelynek eloszlása stabil. Függetlenségi modellje definíció szerint tartalmazza a következőket i=1, . . . , n: $(X_i \perp\!\!\!\perp \{X_1, \ldots X_{i-2}\}|X_{i-1})$, és az implikált $(X_i \perp\!\!\!\perp \{X_1, \ldots X_{i-2}, X_{i+2}, \ldots, X_n\}|\{X_{i-1}, X_{i+1}\})$. Ez a függetlenségi modell n darab azonos lineáris vázú Bayes-hálóval is egzakt módon reprezentálható (azok perfekt térképei), amelyekben nincs összetartó élpár. (Két speciális eset az „előrefele" és a „visszafele" irányított hálózat, lásd 3.2 ábra.)



3.2. ábra. Azon három X_1, X_3 változós Bayes-hálók ekvivalencia-osztályai, amelyekben direkt függés van X_1, X_2 és X_2, X_3 között, de nincs X_1, X_3 között.

A DAG-okból d-szeparációval indukált függetlenségi modellek lehetővé teszik egy DAG-ok feletti ekvivalencia-reláció bevezetését [21, 22, 31].

20. definíció. Két DAG G_1, G_2 *megfigyelési ekvivalens*, ha pontosan ugyanazokat a d-szeparációs relációkat definiálják, azaz $((X \perp\!\!\!\perp Y|Z)_{G_1}) \Leftrightarrow (X \perp\!\!\!\perp Y|Z)_{G_2}$.

Az így definiált ekvivalencia-osztályok igen eltérő számú DAG-ot tartalmaznak. A függések teljes hiányához tartozó ekvivalencia-osztály $n!$ számú DAG-ot tartalmaz, amelyek mindegyike egy változósorrend mentén tartalmaz minden élt. Ezzel szemben az teljes függetlenséghez tartozó ekvivalencia-osztály 1 darab DAG-ot tartalmaz, az üres gráfot. Experimentális vizsgálatok azt jelzik, hogy egy ekvivalencia-osztályba átlagosan 3 DAG tartozik [18]. A nagyságrendek érzékeltetése végett jelezzük, hogy a DAG-ok számosságára csak rekurzív képlet létezik [6]:

$$f(n) = \sum_{i=1}^{n} (-1)^{i+1} \binom{n}{i} 2^{i(n-1)} f(n-i) \; with \; f(0) = 1, \tag{3.5}$$

amelyre felső korlátot jelent n csomópontnál $2^{n(n-1)}$ élkombináció (a DAG-ság kényszere miatt ez kisebb). Ez azonban még akkor is szuper-exponenciális, ha a maximális szülőszámot k-ban maximáljuk. (Vegyük észre, hogy a lehetséges szülői halmazok száma egy adott változó-sorrend esetén is nagyságrendileg n^{kn}, azaz $2^{\mathcal{O}(kn \log n)}$ [11].) A sorrendek, a DAG-ok, és egy adott sorrenddel kompatibilis, szülőszámban korlátos DAG-ok számát a 3.1 táblázat mutatja.

3.1. táblázat. A sorrendek, a DAG-ok, és egy adott sorrenddel kompatibilis, szülőszámban korlátos DAG-ok száma. Az oszlopok sorban a következőket tartalmazzák: a változók számát (n), DAG-ok számát ($|DAG(n)|$), egy sorrenddel kompatibilis DAG-ok számát ($|G_\prec|$), egy sorrenddel kompatibilis DAG-ok számát maximulásan 4 szülőszámmal ($|G_\prec^{|\pi|\leq 4}|$), illetve 2-vel ($|G_\prec^{|\pi|\leq 2}|$), a sorrendek (permutációk) számát ($|\prec|$) a szülői halmazok összszámát sorrend kompatibilis DAG-okban $|\pi^\prec|$ és DAG-okban ha a maximális szülőszám 4 ($||\pi^\prec| \leq 4|$), illetve 2 ($||\pi^\prec| \leq 2|$).

n	$\|DAG(n)\|$	$\|G_\prec\|$	$\|G_\prec^{\|\pi\|\leq 4}\|$	$\|G_\prec^{\|\pi\|\leq 2}\|$	$\|\prec\|$	$\|\pi^\prec\|$	$\|\|\pi^\prec\| \leq 4\|$	$\|\|\pi^\prec\| \leq 2\|$
5	2.9e+004	1e+003	1e+003	6.2e+002	1.2e+002	30	30	24
6	3.8e+006	3.3e+004	3.2e+004	9.9e+003	7.2e+002	62	61	40
7	1.1e+009	2.1e+006	1.8e+006	2.2e+005	5e+003	1.3e+002	1.2e+002	62
8	7.8e+011	2.7e+008	1.8e+008	6.3e+006	4e+004	2.5e+002	2.2e+002	91
9	1.2e+015	6.9e+010	2.9e+010	2.3e+008	3.6e+005	5.1e+002	3.8e+002	1.3e+002
10	4.2e+018	3.5e+013	7.5e+012	1.1e+010	3.6e+006	1e+003	6.4e+002	1.7e+002
15	2.4e+041	4.1e+031	2.1e+027	3.1e+019	1.3e+012	3.3e+004	4.9e+003	5.7e+002
35	2.1e+213	1.3e+179	1.8e+109	8.5e+068	1e+040	3.4e+010	3.8e+005	7.2e+003

Az azonos ekvivalencia-osztályba tartozó DAG-ok tulajdonságainak megértése több szempontból is fontos. Egyrészt szükséges tisztázni a DAG-ok szándékolt, intuitív oksági szemantikájának fenntarthatóságát, nevezetesen azt, hogy milyen korlátok között maradhatna érvényes ez az oksági értelmezés (mint a Markov-láncok 3.2.1 példája mutatja bizonyos esetekben az éleknek semmilyen irányítást nem tulajdoníthatunk). Másrészt azonos megfigyelési ekvivalencia-osztályba tartozó DAG-ok Bayes-hálóit azonos módon kellene felparaméterezni, ami akauzális megközelítésben is fontos következményekhez fog vezetni.
Az azonos ekvivalencia-osztályba tartozó DAG-ok jellemzése két észrevételen nyugszik. Az első, hogy az azonos megfigyelési ekvivalencia-osztályba tartozó DAG-ok irányítatlan váza azonos, mivel a DAG-ban egy él egy közvetlen függést reprezentál, amelynek minden Markov kompatibilis DAG-ban meg kell jelennie [22]. A második észrevétel, hogy ha X, Y és Y, Z közötti közvetlen függések léteznek, úgy, hogy nincs közvetlen függés X, Z között és nincs olyan függetlenség, hogy $(X \perp\!\!\!\perp Z | \{Y, S\})$, azt mindenképpen egy összetartó élpárral kell jelezni $X \rightarrow Y \leftarrow Z$, egy úgynevezett *v-struktúrát* létrehozva.
Az azonos ekvivalencia-osztályba tartozó DAG-ok jellemzését a következő tétel biztosítja.

3.2.1. Tétel. [[5, 22]] Két DAG G_1, G_2 pontosan akkor *megfigyelési ekvivalens*, ha az irányítatlan vázuk megegyezik és ugyanazon *v-struktúrákat* tartalmazzák (azaz konvergáló éleket, amelyek talpánál nincs él) [22]. . Ha a Bayes-hálók $(G_1, \underline{\theta}_1)$ és $(G_2, \underline{\theta}_2)$ diszkrét változókat tartalmaznak és lokális modelljeik multinomiális eloszlások, akkor G_1, G_2 megfigyelési ekvivalenciája egyenlő dimenzionalitást és bijektív leképezhetőséget jelent a $\underline{\theta}_1$ és $\underline{\theta}_2$ paraméterezések között, amit *eloszlásbeli ekvivalenciának*neveznek [5]).

Mint látható, ha elfogadjuk az Ockham-elv által diktált modellminimalitás elvét, és egy eloszlásmodellezésnél (az egyszerűség kedvéért stabil eloszlást feltételezve) a függetlenségi modelljét minimális módon reprezentáló DAG-okat tekintjük, akkor bizonyos élek irányítása önkényes, így oksági értelmezése, a priori információk hiányában értelmetlen. Azonban a 3.2.1 tételben szereplő v-struktúráknál több élre jelenthet megkötést a megfigyelési osztályba tartozás, hiszen bizonyos élek irányítása azért lehet egyértelmű, mert amúgy v-struktúrát hoznának létre (ami kivezetne az ekvivalencia-osztályból). Ez a következő definícióhoz vezet el.

21. definíció. Az *esszenciális gráf* a megfigyelési ekvivalens DAG-ok halmazát reprezentálja egy részlegesen irányított DAG-gal (PDAG), amely gráfban csak azok az úgynevezett kényszerített élek irányítottak, amelyek az ekvivalenciaosztálybeli DAG-okban azonosan irányítottak. A többi él irányítatlansága az (élszintű) eldönthetetlenséget jelzi.

Az esszenciális gráf meghatározására hatékony algoritmust közölt Meek [21].

3.3. Oksági Bayes-hálók

A klasszikus kérdés, hogy hogyan lehet megkülönböztetni az oksági kapcsolatokat a függésektől („korreláció versus kauzalitás"), azaz, hogy hogyan lehetne meghatározni az

oksági státuszát passzívan megfigyelt X és Y közötti statisztikai függésnek, az felbontható a valószínűségi Bayes-hálós reprezentációkhoz tartozó fogalmakkal, mint stabilitás és az esszenciális gráf. Elsőként megfontolandó, hogy vajon az összes közvetlen függés oksági-e. Ez erősen vitatható feltevés volna, amelyre hosszabban kitérünk. Másodsorban a stabilitás feltevése is megfontolható, hiszen annak hiányában (a Bayes-hálós reprezentáció definíciója szerint) nem fennálló függéseketis impliklni fog a struktúra. Harmadsorban, meg lehet fontolni, hogy az esszenciális gráf és a kényszerített élek definiálásánál használt „Boolean" Ockham elv (amely szerint csak a minimális, konzisztens modelleket vettük figyelembe) a bayesi kontextusban nem terjeszthető-e ki?
Ezen kérdések megfontolásához vezessük be az oksági modell fogalmát, amely a korábbi, Bayes-hálókon alapuló intuíciót formalizálja.

22. definíció. Egy DAG-ot a *oksági struktúrának* nevezünk változók V halmaza felett, ha minden csomópont egy változót reprezentál, az élek pedig közvetlen ráhatást szimbolizálnak. Egy *oksági modell* olyan oksági struktúra *lokális valószínűségi modellekkel* $p(X_i|\,\mathrm{pa}(X_i))$ minden egyes csomóponthoz, amely leírja az adott X_i csomópont sztochasztikus függését a $\mathrm{pa}(X_i)$ szüleitől. Mivel a feltételes modellek gyakran parametrikus modellcsaládból származnak, az X_i-hez tartozó feltételes modell paramétereit $\underline{\theta}_i$ jelöli, és $\underline{\theta}$ jelöli a teljes modell paraméterezését.

A stabilitás feltevésével az esszenciális gráf egzakt módon reprezentálja a függetlenségi relációkat, és a Boolean Ockham elv szerinti modellminimalitásnak megfelelően maximális mértékben jelzi a potenciális oksági relációkat, így elfogadásával az oksági relációk rendszer–alapú kikövetkeztetésére láthatnánk példát. A feltevések jogosságának vizsgálatához vezessük be az alábbi formális feltételt, amely egy oksági struktúra validitását és elégségességét biztosítja.

23. definíció. Egy G oksági struktúra és p eloszlás teljesíti az *oksági Markov-feltételt (CMA*, ha p-ben teljesül a G szerinti lokális Markov-feltétel.

Az oksági Markov-feltétel Reichenbach „közös ok elv"-én alapul, amely szerint X és Y események közötti függés azért áll fenn, mert vagy X okozza Y-t, vagy Y okozza X-t, vagy közös ok befolyásolja X-t és Y-t is [14, 23]. Ennek megfelelően az oksági Markov-feltétel akkor áll fenn (p, G) párra, ha a V változóhalmaz *okságilag elégségséges*, azaz nincs rejtett, nem V-beli, közös ok (vagy másképpen fogalmazva: minden közös ok $X, Y \in V$ párokra V-beli). Ez természetesen nem azt jelenti, hogy nem lehetnek rejtett változók, hiszen ez egy adott absztrakciós szinten elkerülhetetlen, de csak azon változóknak szükséges V-ben szerepelni, amelyek két vagy több változót is közvetlenül befolyásolnak.
Az oksági Markov-feltétel összekapcsolja az oksági relációkat és a függéseket, és az oksági modell (modellezés) elégségességét követeli meg a megfigyelt függésekhez (mondhatni úgy is, hogy az élek elégségesek). Érdemes észrevenni, hogy a stabilitás feltevése éppen az élek szükségességét jelenti (mondhatni úgy is, hogy nincsen felesleges él). Ez a két feltevés biztosíthatja, hogy a Bayes-háló által implikált függetlenségek valóban fennállnak és a függések is egzakt módon reprezentáltak az oksági modellben [12]. Az oksági Markov-feltétel lehetővé teszi továbbá a beavatozások modellezését $do()$ művelet (19) bevezetésével a „*manipulációs tétel*" ([28]) avagy „gráf csonkolás" ([23]) szerint.

24. definíció. Egy $G, \underline{\theta}$ oksági modell esetén $p(Y|z, do(X = x))$ jelölje azt az eloszlást, amelyet úgy kapunk, hogy a (perfekt) beavatkozáshoz tartozó X változók bemenő éleit töröljük és ezeket a változókat az előírt értékre beállítjuk (azaz a következtetés során a beállított változókhoz tartozó faktorok nem szerepelnek).

A beavatkozások formális modellje jelzi, hogy egy oksági modell, amely teljesíti az oksági Markov-feltételt, minden lehetséges (perfekt) beavatkozáshoz tartozó eloszlást a fenti gráf csonkolásos szemantikával képes reprezentálni [23]. Az oksági modellek és a beavatkozások kapcsolata tovább is vihető, ami elvezet az autonóm, lokális „mechanizmusok" rendszeréhez, amelyek beavatkozásokra függetlenül reagálnak, és zajjal való terheltségük is független. A „*funkcionális Bayes-hálók*" ezen zajjal terhelt determinisztikus mechanizmusokon alapulnak, kapcsolódva a strukturális egyenletek formalizmusához is [9, 23]. A funkcionális Bayes-hálók a beavatkozásokon kívüli kontrafaktuális következtetésekben is felhasználhatóak, ami egy elképzelt múltbeli beavatkozás jelen következményeinek következtetését jelenti.

3.4. Az oksági értelmezés nehézségei

Az oksági értelmezés kritikája előtt érdemes összefoglalni a valószínűségi modellezés számára is kihívást jelentő lehetõségeket.

3.4.1. Tisztán magasabbrendű függések

Tegyük fel, hogy X, Y, Z bináris változók, X, Y függetlenek, egyenletes eloszlásúak, Z pedig a $Z = XOR(X, Y)$ logikai függvénnyel meghatározott. Ekkor $(X \perp\!\!\!\perp Z)$ és $(Y \perp\!\!\!\perp Z)$, de $(\{X, Y\} \not\perp\!\!\!\perp Z)$, azaz az együttesüktől már függ. A függések (asszociációk) tehát nem feltétlenül monotonak.

Függések iránya

Egy Markov-lánc $\mathcal{X} = \{X_0, X_1, \ldots X_n\}$, amely a $p(X_i|X_{i-1})$ átmeneti valószínûségekkel adott, átírható $p(X_{i-1}|X_i)$ alakba is, azaz a lánc irányítása esetleges.

3.4.2. Intranzitív függések

Tegyük fel, hogy X, Y, Z közül legalább egy nem bináris változó, például Y. Ekkor létezik egy olyan $p(X, Y, Z)$ eloszlás, amelyre $(X \not\perp\!\!\!\perp Y)$ és $(Y \not\perp\!\!\!\perp Z)$, de $(X \perp\!\!\!\perp Z)$, azaz a függések (asszociációk) nem feltétlenül tranzítívak. Azonban ha feltesszük, hogy az eloszlásunk függetlenségi viszonyai stabilak, azaz a kvalitív függési viszonyok nem változnak infinitizimális perturbációkra, akkor az intranzitív hármas függési modellt egyedül az úgynevezett v-struktúra magyarázhatja $(X \to Y \leftarrow Z)$, amelyben $(X \not\perp\!\!\!\perp Z|Y)$ (és nem $(X \perp\!\!\!\perp Z|Y)$). Azaz ekkor kivételesen adódik, hogy Y a következménye két őt befolyásoló független eseménynek, amelyek tapasztalataink szerint mindig korábbiak (innen ered a filozófiában „idõ nyila" elnevezés).

3.4.3. Simpson paradoxona

Lehetséges olyan eloszlás, amelyben $p(Y|X) > p(Y|\bar{X})$, de $p(Y|X,Z) < p(Y|\bar{X},Z)$ és $p(Y|X,\bar{Z}) < p(Y|\bar{X},\bar{Z})$, azaz az asszociáció hatása rétegenként (Z szerint) és összesítve ellenkező is lehet.

3.4.4. Ellenérvek

Az oksági értelmezéssel szemben felhozott ellenérvek egy része a fenti, a valószínűségi modellezést is érintő kérdésekhez kapcsolódik, más része a magának a Bayes-hálós megközelítésnek a problémájára világít rá (például az okság fogalma visszacsatolt vagy emergens rendszerekben). Mivel ezek tárgyalása meghaladja a jegyzet kereteit, csupán felsorolásban jelezzük őket [14, 28].

1. A direkt függések nem mindegyike oksági (hanem például szemantikai).
2. Adott eloszlás esetén a hozzátartozó megfigyelési ekvivalencia-osztály egyértelműségéhez fel kell tételezni a stabilitást (ami logikai függéseknél nem teljesül, ahogyan azt a XOR-t tartalmazó példa mutatta).
3. Az esszenciális gráf definíciója csak a minimális, konzisztens modellekre támaszkodik (egyáltalán nem véve figyelembe a kissé komplexebb konzisztens modellek irányítását).
4. Rejtett zavaró tényezők, stabilitás, modell minimalitás.
5. Kiválasztási bias, amikor is a megfigyelés akár független események kombinációjától függ, így okozva az adatokban akauzális függést.
6. Oksági modellek keveréke, azaz ha X befolyásolja Y-t és fordítva. Hasonló probléma a visszacsatolás.
7. Globális fizikai és szemantikai kényszerek a változók között.
8. Az asszociációk, a függetlenségek, a függések és így az induktívan kikövetkeztetett oksági viszonyok viszonylagosak az elemzett változók halmazához képest, sőt még az értéktartományukhoz (diszkretizálásukhoz) képest is.

Mindazonáltal kijelenthető, hogy a többváltozós elemzések miatt és a Bayes-statisztikai megközelítés miatt az oksági következtetést ma már széles körben alkalmazzák, meghaladva a korábbi két-változós „correlation (association) $\neq$ causation" zsákutcás igaz/-hamis szemléletet.

3.5. Bayes-hálók a Bayes-statisztikai keretben

A nagy dimenziós, viszonylagosan kis mintás orvosbiológiai alkalmazások az egyik fő terepe a Bayes-hálók Bayes-statisztikai keretrendszerben való alkalmazásának. Természetesen a modellosztály választása a statisztikai kerettől független dimenzió, de az oksági modellek feletti átlagolás igénye a többi modellosztály használatával egyidőben

felvetődött [8]. Ez megjelent a valódi ok kényszerített éleken alapuló bayesi megközelítésében (lásd például [10, 17]) vagy a hatáserősség struktúrák feletti átlagolást is magába foglaló becslésében [23]), illetve az optimális intervención alapuló adatgyűjtésben [34].
Azonban a Bayes-statisztikai megközelítés alkalmazása Bayes-hálókra rengeteg megoldatlan kérdést vet fel a modellhez kapcsolódó a priori ismeretek sokfélesége miatt. Ez természetesen a modell értelmezésének sokféleségével is összefügg, azonban ezek egységes kvantitatív formába történő transzformálása, azaz az őket tükröző informatív a priori eloszlások konstruálása jelenleg is aktív kutatás alatt áll, részben a „transfer" learning részeként is. Logikai kényszerek felhasználása [4, 6, 19, 29], kvalitatív monotonitási relációk felhasználása [15, 30] sok egyéb mellett már megjelent tudományos közleményekben mint prior eloszlások információforrása [1, 2, 6, 16, 26].
A Bayes-statisztikai megközelítésben a prior eloszlás $p(G, \underline{\theta})$) a DAG-struktúrák és hozzájuk tartozó paraméterek felett definiált. A Bayes-háló reprezentáció univerzalitása miatt ez természetesen az eredeti tárgyterület feletti Bayes-háló kiterjesztésében is reprezentálható. Azonban ennek az együttes a priori eloszlásnak a specifikálása vagy akár csak a paraméterekre vonatkozó $p(\underline{\theta}|G)$ feltételes eloszlásnak a specifikálása, mind elméleti, mind gyakorlati megfontolásokat igényel, egyrészt a modelltér komplex és nagydimenziós volta, másrészt a struktúra- és paraméterekvivalenciák miatt. Elsőként a paraméterpriorokat tárgyaljuk, majd ezt követően a struktúrák feletti priorok kérdését.

3.5.1. Paraméter priorok Bayes-hálókhoz

A $p(\underline{\theta}|G)$ paraméter prior specifikálása a következő kérdéseket veti fel: milyen eloszláscsaládot használjunk, mi a kapcsolata a prior dekomponálásának és a tárgyterületi eloszlás dekomponálásának, hogyan lehet konzisztens módon bizonyosságot definiálni a dekomponált prior felett az egész struktúrát tekintve, hogyan lehet konzisztens priort definiálni a megfigyelési ekvivalens struktúrák tekintetében. Ezeket a kérdéseket egy lentebb, több lépésben kifejtett, átfogó eredmény válaszolta meg, amely nem csak az oksági modellek Bayes-statisztikai felhasználásához, hanem a Bayes-hálók valószínűségi, akauzális felhasználásához is szükséges. Első része kimondja, hogy ha a paraméterprior a struktúra szerint dekomponálódik és a paraméterpriorok ekvivalensek megfigyelési ekvivalens struktúrák esetében, akkor a paraméterprior szükségszerűen Dirichlet eloszlású. Továbbá ha a dekomponált paraméterprior részei a struktúrára nézve invariánsak, akkor a $p(\underline{\theta}|G)$ paraméterprior tetszőleges G struktúra esetén egyetlen pont-értékű θ_0 paraméterezésből és egyetlen a priori mintaszámként értelmezhető skalárból származtatható. Ennek a formális kimondásához és megértéséhez a következő fogalmakra van szükség. Elsőként a paraméter függetlenségre [7, 27]:

25. definíció. Egy G Bayes-háló struktúra esetén, a *globális paraméter függetlenség* feltevése azt jelenti, hogy

$$p(\underline{\theta}|G) = \prod_{i=1}^{n} p(\underline{\theta}_i|G), \tag{3.6}$$

ahol $\underline{\theta}_i$ jelöli a paramétereket, amelyek $p(X_i|\operatorname{Pa}(X_i))$ feltételes eloszláshoz tartoznak

G-ben. A *lokális paraméter függetlenség* feltevése azt jelenti, hogy

$$p(\underline{\theta}_i|G) = \prod_{j=1}^{q_i} p(\underline{\theta}_{ij}|G), \tag{3.7}$$

ahol q_i a szülői konfigurációk számát jelenti ($\mathrm{pa}(X_i)$) X_i-hez tartozik G-ben és $\underline{\theta}_{ij}$ a paramétereket jelöli a $p(X_i|\mathrm{pa}(X_i)_j)$ feltételes eloszlásban valamely fix sorrendjében a $\mathrm{pa}(X_i)$ konfigurációknak. A *paraméter függetlenség* feltevése mind a globális, mind a lokális függetlenség feltevését jelenti.

A likelihood ekvivalencia fogalma a megfigyelési ekvivalenciát terjeszti ki a struktúrákról a paraméterekre ([13, 16]).

26. definíció. A *likelihood ekvivalencia* feltevés azt jelenti, hogy két megfigyelési ekvivalens Bayes-háló struktúra G_1, G_2,

$$p(\underline{\theta}_V|G_1) = p(\underline{\theta}_V|G_2), \tag{3.8}$$

ahol $\underline{\theta}_V$ a multinomiális parametéreknek egy nem redundáns halmazát jelenti a teljes V együttes eloszlásra nézve. (A lokális modellek multinomiális volta biztosítja az eloszlás ekvivalenciát és, azt, hogy Jacobi-paramétertranszformáció létezik.)

Ezek után a következő tétel mondható ki [13, 16].

3.5.1. Tétel. [[13, 16]] Pozitív sűrűségfüggvények, likelihood ekvivalencia és paraméter függetlenség feltevése G_c teljes struktúrákra azt implikálja, hogy $p(\underline{\theta}_V)$ szükségszerűen Dirichlet eloszlású $N_{x_1,\ldots,x_n}$ hiperparaméterekkel.

A $p(\underline{\theta}_i|G_i) = J_{G_i}p(\underline{\theta}_V)$, ahol J_{G_i} a Jacobi transzformáció $\underline{\theta}_V$-ről $\underline{\theta}_{G_i}$-re. Figyelemre méltó, hogy egy struktúrák szintjén megfogalmazott kényszer, a struktúrák likelihood ekvivalenciája multinomiális lokális modellekkel, ilyen erős paraméterszintű kényszert eredményez. A következő eredmény kimondásához írjuk át a hiperparamétereket, mint $N' = \sum_{x_1,\ldots,x_n} N_{x_1,\ldots,x_n}$, amit *prior vagy virtuális mintaméretnek* hívnak és $p(x_1, \ldots, x_n|\xi^+) = N_{x_1,\ldots,x_n}/N'$. Továbbá, szükséges még a következő fogalom:

27. definíció. A *paraméter modularitás* feltevése azt jelenti, hogy ha $\mathrm{pa}(X_i)$ azonosak két Bayes-háló struktúrában G_1, G_2, akkor

$$p(\underline{\theta}_{ij}|G_1) = p(\underline{\theta}_{ij}|G_2), \tag{3.9}$$

ahol $\underline{\theta}_{ij}$ jelöli a paramétereket a kapcsolódó $p(X_i|\mathrm{pa}(X_i)_j)$ feltételes eloszlásban a $\mathrm{pa}(X_i)$ konfigurációk valamely fix sorrendje esetében.

Az oksági értelmezéshez közel álló paraméter modularitás feltevése lehetővé teszi paraméterprior származtatását a teljes modellekről nem teljes modellekre is .

3.5.2. Tétel. [[13, 16]] Ha N' a globális prior mintaméret, $p(\underline{\theta}_V)$ egy Dirichlet eloszlás $N_{x_1,\ldots,x_n} = N'p(x_1, \ldots, x_n)$ hiperparaméterekkel, továbbá feltesszük a paraméter modularitást, és minden G_c teljes DAG-ra, $p(G_c) > 0$, akkor bármely G struktúrára teljesül

a paraméter függetlenség és likelihood ekvivalencia és a paraméterek dekomponált eloszlása Dirichlet eloszlásoknak a következő szorzata:

$$p(\underline{\theta}|G,\xi^+) \;=\; \prod_{i=1}^{n}\prod_{j=1}^{q_i}\prod_{k=1}^{r_i}\theta_{ijk}^{N'p(X_i=k,\mathrm{pa}(X_i,G|\xi^+)=\mathrm{pa}_{ij})-1}, \tag{3.10}$$

ahol r_i az X_i változó értékeinek száma, q_i a $\mathrm{pa}(X_i,G)$ szülői halmaz lehetséges értékkonfigurációinak száma, és pa_{ij} jelöli a szülők értékeit a j-edik szülői konfigurációban, valamely fix sorrendjében a szülői konfigurációknak.

A 3.5.2 tétel praktikus módszert kínál likelihood ekvivalens paraméterpriorok megadására minden struktúrára: egy maximálisan részletes modell esetében határozzuk meg a pontparametrizációt és az a priori mintaszámot, majd bármely más modell esetén marginalizáljuk az eloszlást, és számítsuk ki az ott releváns hiperparamétereket. Azonban a 3.5.2 tétel azt is jelzi, hogy nem teljes megfigyelések esetében a struktúra különböző részein eltérő bizonyosság fog jelentkezni, így nem lehetséges egyetlen mintaszámmal jellemezni azt [14].

3.5.2. Struktúra priorok Bayes-hálókhoz

A bayesi megközelítés Bayes-hálós modellek paramétereihez az 1980-as évektől jelen van a szakirodalomban [7, 25, 26], és ez a kutatási irány részben megválaszolta a komplex valószínűségi modellek paraméterezésével kapcsolatos ellenvetéseket [3]. A struktúrák feletti bayesi megközelítés az 1990-es évek elején jelent meg, de a nagy számításigény sokáig gátolta az alkalmazását. Egy sorrendspecifikus, analitikus megközelítés [1], majd egy általános analitikai eredmény [6], ezt követően pedig 1995-ben, az MCMC-módszerek alkalmazása is megjelent [20]. 2000. óta a modellek feletti bizonytalanság kezelésére elterjedtté vált a struktúrák feletti bayesi megközelítés alkalmazása, azonban struktúrális háttérinformációk felhasználása a mai napig nem megoldott. Vegyük észre, hogy a $p(G)$ struktúra priorok kiegészítik a korábbiakban tárgyalt $p(\underline{\theta}|G)$ paraméterpriorok tárgyalását.

Referencia struktúrák és alstruktúrák felhasználása Az egyik alapvető módszer, az *eltérés alapú* priorok egy referencia struktúrától való eltérést büntetnek, felhasználva egy „referencia" struktúrát G_0 és egy κ büntető faktort a hiányzó vagy extra e_{ij} élek büntetésére [16]:

$$p(G) \propto \kappa^{\delta}, \text{ahol } \delta = \sum_{1\le i<j\le n} 1(1(e_{ij}\in G) \neq 1(e_{ij}\in G_0)).$$

A *jegy alapú priorok* az egyes jegyek jelenléte szerint definiáltak, ahol az F_i jegy (modelltulajdonság) G-beli értékeit $F_i(G) = f_i$ jelöli, $i = 1, \ldots K$ esetén,

$$p(G) = c\prod_{i=1}^{K} p(F_i(G)), \tag{3.11}$$

ahol a c normalizációs konstans az inkonzisztens jegykombinációk kezelésére szolgál. Lehetséges jegyek az irányítatlan vagy irányított vagy kényszerített élek, páronkénti

vagy részleges sorrend szerinti sorrendezések fennállása, Markov-határbeliség vagy akár tetszőleges algráf megléte. Érdemes ügyelni a jegyek összefüggő voltára, ami a globális DAG-kényszer miatt lép fel, és torzítja az egyenletben szándékolt független hatást.

3.6. Megfigyelés, beavatkozás, spekuláció

A valószínűségi, oksági, és funkcionális Bayes-hálók a következtetés különböző szintjeit képesek támogatni, amit csupán felsorolásszerűen összegzünk (kifejtését lásd [23]):

1. Jelölje $p(Y = y|X = x)$ a (megszokott megfigyelési) feltételes valószínűségét $Y = y$ értéknek $X = x$ érték megfigyelése esetén.

2. Jelölje $do(x)$ az X változó x értékre történő beállításához tartozó beavatkozást és $p(Y|do(x))$ az ehhez tartozó beavatkozási eloszlást.

3. Jelölje $p(Y = y|do(X = x), Y = y', X = x')$ a kontrafaktuális valószínűségét annak, hogy $Y = y$, amikor $X = x', Y = y'$ és $do(X = x)$.

3.7. Tudásmérnökség

A Bayes-statisztikai keretben definiált Bayes-háló a tudásmérnökség eszközeként jelent meg a 1980-as években. Konstruálása jellemzően a szakértőktől származó adatokból történt manuálisan. A kézi konstruálás még napjainkban is jelentős súlyt képvisel a Bayes-hálók alkalmazásában. Azonban ahol az adathoz viszonyítva jelentős mennyiségű a priori tudás áll rendelkezésre, ott a Bayes-hálók tudásmérnöki alkalmazása prior konstruálás formájában is előfordul, a bayesi keretrendszer alkalmazásának egy kezdeti fázisaként. A tudásmérnökség metodikájára nagy hatással volt a nagy mennyiségű elektronikus tárgyterületi információ megjelenése, a megfelelő mennyiségű statisztikai adat elérhetősége, valamint a Bayes-statisztikai alapú gépi tanulási módszerek elterjedése. A modell konstruálása helyett érdemes a gyakorlatban tipikusan megjelenő egész tudásbázist tekintetbe venni, és annak konstruálására fókuszálni. Követelményként jelent meg a bayesi módszerek alkalmazásakor, hogy támogassa a priorok konstruálását, hiszen a valószínűségekkel leírt a priori tudás és a rendelkezésre álló adatok bayesi frissítéssel történő kombinációja szolgáltatja a végső tudásmodellt. Mindemellett fontos, hogy a tudásbázis segítse a komplex, akár szabad szöveges háttérismereteket is tartalmazó valószínűségi állítások megfogalmazását, valamint tegye lehetővé a szakértőktől származó szubjektív információ tárolását, mely releváns lehet a bayesi a priori tudásmodell megalkotásánál. Egy tudásbázis megépítéséhez olyan környezetben, ahol rendelkezésre áll elektronikus tárgyterületi tudás, elegendő statisztikai adat, valamint a megfelelő bayesi módszerek, az alábbi lépések szükségesek (amelyekből a specifikusokat részletezzük):

1. Célok, alkalmazási terület és modellezési szintek identifikációja. Terminológia és ontológia elfogadása.

2. Nem rendszerezett tudás begyűjtése. Ehhez a lépéshez tartozik az összes releváns elektronikus és egyéb szövegalapú információforrás feldolgozása, amely magába

foglalja az a priori információ kinyerését különféle szövegbányászati metódusok alkalmazásával.

3. Struktúra kinyerése. A G DAG struktúrák feletti p(G) priorok konstruálása, melyek egyesítik a szakértők által megadott információkat az elektronikus forrásokból kinyert információkkal (a p(G) a priori eloszlást többnyire normalizálatlan formában lehet előállítani).

4. Paraméter és hiperparaméter kinyerése. A valószínűségi paraméterek többféle módon nyerhetők: adatbázisok, szakirodalom vagy szakértők szubjektív véleménye alapján. A $p(\theta|G)$ paraméterprior specifikációja az általunk vizsgált diszkrét, véges esetben egy egyszerű módszerrel megvalósítható, ha feltehetjük az egyes változókhoz és szülői értékkonfigurációkhoz tartozó paraméterek függetlenségét. Egy szinte kizárólagosan használt eloszláscsalád az adott változó, adott szülői értékkonfigurációjához tartozó feltételes modellek megadására a Dirichlet eloszlás, amelyben a hiperparaméter a paraméterhez tartozó szülői értékkonfiguráció korábban megfigyelt eseteinek számait jelenti (Cowel1999). Ahogyan megmutattuk a 3.5.1 tételben, a Dirichlet család az egyetlen lehetséges választás, ha az ugyanazon megfigyelési ekvivalencia-osztályba tartozó G struktúrákhoz ekvivalens priorokat szeretnénk megadni, ami kauzális modellezésnél nem szükségszerű (Heckerman1995a).

5. Érzékenységi analízis, verifikáció és validáció. A modellek poszteriorjának vizsgálata magába foglalja egyrészt az a priori eloszlásokra való érzékenység vizsgálatát (ami különösen fontos a több szakértőt és tudásbázist is felölelő automatizáltan származtatott prioroknál), másrészt referencia priorokkal való összehasonlítást. A modellosztály komplexitása miatt mindkét esetben gyakran szükséges, hogy egyrészt modell jegyeket használjunk, másrészt hogy MAP-modellre alapozzuk a vizsgálatot.

Mint ahogy az látható, a tudásbázis építése a bayesi modellkiértékeléssel, modelfinomítással, esetleg tanulással zárul. A kiértékelés tartalmazza az adat és a modell kompatibilitásának vizsgálatát és az a posteriori valószínűségek vizsgálatát, más esetekben a tudásmérnöki folyamat célja az a priori modell konstruálása a későbbi tanulási folyamat számára.

3.8. Bayes-háló kiterjesztések

A beavatkozásokkal történő következtetésekre formális lehetőséget kínálnak a döntési hálók, amelyekben a véletlen csomópontok mellett beavatkozás- és hasznosságcsomópontok is találhatóak. Használatukat és a valószínűségi, illetve az oksági Bayes-hálók egyéb kiterjesztéseit a Szekvenciális döntéstámogatás című fejezetben, illetve Bioinformatikai laboratórium jegyzet fejezeteiben tárgyaljuk.

Irodalomjegyzék

[1] W. L. Buntine. Theory refinement of Bayesian networks. In *Proc. of the 7th Conf. on Uncertainty in Artificial Intelligence (UAI-1991)*, pages 52–60. Morgan Kaufmann, 1991.

[2] R. Castelo and A. Siebes. Priors on network structures. biasing the search for Bayesian networks. *International Journal of Approximate Reasoning*, 24(1):39–57, 2000.

[3] P. Cheeseman. In defense of probability. In *Proceedings of the Ninth International Joint Conference on Artificial Intelligence (IJCAI-85)*, pages 1002–1009. Morgan Kaufmann, 1985.

[4] J. Cheng, D. A. Bell, and W. Liu. Learning belief networks from data: an information theory based approach. In *Proc. of the 6th ACM International Conference on Information and Knowledge Management, CIKM'97*, pages 325–331, 1997.

[5] D. M. Chickering. A transformational characterization of equivalent Bayesian network structures. In *Proc. of 11th Conference on Uncertainty in Artificial Intelligence (UAI-1995)*, pages 87–98. Morgan Kaufmann, 1995.

[6] G. F. Cooper and E. Herskovits. A Bayesian method for the induction of probabilistic networks from data. *Machine Learning*, 9:309–347, 1992.

[7] R. G. Cowell, A. P. Dawid, S. L. Lauritzen, and D. J. Spiegelhalter. *Probabilistic networks and expert systems*. Springer-Verlag, New York, 1999.

[8] A. P. Dawid. Discussion of 'causal diagrams for empirical research' by j. pearl. *Biometrika*, 82(4):689–690, 1995.

[9] M. J. Druzdzel and H. Simon. Causality in Bayesian belief networks. In David Heckerman and Abe Mamdani, editors, *Proceedings of the 9th Conf. on Uncertainty in Artificial Intelligence (UAI-1993)*, pages 3–11. Morgan Kaufmann, 1993.

[10] N. Friedman, M. Goldszmidt, and A. Wyner. On the application of the bootstrap for computing confidence measures on features of induced Bayesian networks. In *AI&STAT VII*, 1999.

[11] N. Friedman and D. Koller. Being Bayesian about network structure. In *Proc. of the 16th Conf. on Uncertainty in Artificial Intelligence(UAI-2000)*, pages 201–211. Morgan Kaufmann, 2000.

[12] D. Galles and J. Pearl. Axioms of causal relevance. *Artificial Intelligence*, 97(1-2):9–43, 1997.

[13] D. Geiger and D. Heckerman. A characterization of the Dirichlet distribution with application to learning Bayesian networks. In Philippe Besnard, Steve Hanks, Philippe Besnard, and Steve Hanks, editors, *Proc. of the 11th Conf. on Uncertainty in Artificial Intelligence (UAI-1995)*, pages 196–207. Morgan Kaufmann, 1995.

[14] C. Glymour and G. F. Cooper. *Computation, Causation, and Discovery.* AAAI Press, 1999.

[15] A. J. Hartemink, D. K. Gifford, T. S. Jaakkola, and R. A. Young. Bayesian methods for elucidating genetic regulatory networks. *IEEE Intelligent Systems*, 17(2):37–43, 2002.

[16] D. Heckerman, D. Geiger, and D. Chickering. Learning Bayesian networks: The combination of knowledge and statistical data. *Machine Learning*, 20:197–243, 1995.

[17] D. Heckermann, C. Meek, and G. Cooper. A Bayesian aproach to causal discovery. Technical Report, MSR-TR-97-05, 1997.

[18] T. Kocka and R. Castelo. Improved learning of Bayesian networks. In Jack S. Breese and Daphne Koller, editors, *Proc. of the 17th Conference on Uncertainty in Artificial Intelligence (UAI-2001)*, pages 269–276. Morgan Kaufmann, 2001.

[19] W. Lam and F. Bacchus. Using causal information and local measures to learn Bayesian networks. In David Heckerman and Abe Mamdani, editors, *Proc. of the 9th Conference on Uncertainty in Artificial Intelligence (UAI-1993)*, pages 243–250. Morgan Kaufmann, 1993.

[20] D. Madigan, S. A. Andersson, M. Perlman, and C. T. Volinsky. Bayesian model averaging and model selection for Markov equivalence classes of acyclic digraphs. *Comm.Statist. Theory Methods*, 25:2493–2520, 1996.

[21] C. Meek. Causal inference and causal explanation with background knowledge. In *Proc. of the 11th Conf. on Uncertainty in Artificial Intelligence (UAI-1995)*, pages 403–410. Morgan Kaufmann, 1995.

[22] J. Pearl. *Probabilistic Reasoning in Intelligent Systems.* Morgan Kaufmann, San Francisco, CA, 1988.

[23] J. Pearl. *Causality: Models, Reasoning, and Inference.* Cambridge University Press, 2000.

[24] A. Rosenberg. *Philosophy of Science: A contemporary introduction.* Routledge, 2000.

[25] D. J. Spiegelhalter. Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society, Series B*, 50(2):157–224, 1988.

[26] D. J. Spiegelhalter, A. Dawid, S. Lauritzen, and R. Cowell. Bayesian analysis in expert systems. *Statistical Science*, 8(3):219–283, 1993.

[27] D. J. Spiegelhalter and S. L. Lauritzen. Sequential updating of conditional probabilities on directed acyclic graphical structures. *Networks*, 20(.):579–605, 1990.

[28] P. Spirtes, C. Glymour, and R. Scheines. *Causation, Prediction, and Search*. MIT Press, 2001.

[29] S. Srinivas, S. Russell, and A. Agogino. Automated construction of sparse Bayesian networks for unstructured probabilistic models and domain information. In *Proc. of the 5th Conference on Uncertainty in Artificial Intelligence (UAI-1990)*, pages 295–308. North-Holland, 1990.

[30] A. Tanay and R. Shamir. Computational expansion of genetic networks. *Proc. of Int. Conf. on Intelligent Systems for Molecular Biology (ISMB'01)*, 17(Suppl. 1):270–278, 2001.

[31] T. Verma and J. Pearl. *Equivalence and synthesis of causal models*, volume 6, pages 255–68. Elsevier, 1990.

[32] J. Woodward. Scientific explanation. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*, 2003.

[33] M. Woodward. *Epidemiology: Study design and data analysis*. Chapman&Hall, 1999.

[34] Changwon Yoo and Gregory F. Cooper. An evaluation of a system that recommends microarray experiments to perform to discover gene-regulation pathways. *Artificial Intelligence in Medicine*, 31:169–182, 2004.

4. fejezet

Tudásmérnökség, biasok és heurisztikák becsléseknél és döntéseknél

4.1. Valószínűségi ítéletalkotás és a bayesi paradigma

A valószínűségi ítéletalkotás lényegi kérdése, hogy milyen módon hasznosítható a rendelkezésre álló információ bizonytalan helyzetekben. A döntést számos tényező befolyásolhatja, melyek egy része objektív, más része szubjektív faktorokban testesül meg. Jelentős különbségek tapasztalhatók az emberi döntési mechanizmusokban többek között aszerint, hogy mennyi idő áll rendelkezésre a döntés meghozatalára, vagy milyen mértékben jósolható meg előre egy adott típusú eseménysor kimenetele. Mindezek mellett olyan tényezők is közrejátszanak, mint a rendelkezésre álló információba vetett bizalom, az adott probléma kezelésébe való bevonódás mértéke (a részletekre való rálátás mértéke), a hasonló szituációk kezelésében szerzett korábbi kudarc- vagy sikerélmények.
Ha pusztán matematikai, azaz normatív szempontból közelítjük meg a valószínűségi ítéletalkotást, akkor a bayesi paradigma alapja, a Bayes-tétel alkalmazása a megfelelő eszköz. Jelölje H_i az i-edik hipotézist, míg D a rendelkezésre álló adatot. A cél annak a feltételes valószínűségnek a megadása, mely szerint H_i igaz D adat esetén.

$$P(H_i|D) = \frac{P(D|H_i) \cdot P(H_i)}{P(D)}, \tag{4.1}$$

ahol $P(D|H_i)$ a likelihood, azaz annak a feltételes valószínűsége, hogy D bekövetkezik, ha H_i igaz, $P(H_i)$ pedig a H_i hipotézis igaz voltának előzetes (a priori) valószínűségét jelöli. A $P(H_i|D)$-t más néven utólagos vagy a posteriori valószínűségnek nevezzük. E tétel alkalmazása lehetőséget ad a kezdeti szubjektív bizonytalanság optimális felülbírálatára új információk alapján. Ezt bizonytalanságrevíziónak nevezzük, amely a valószínűségi következtetés és döntéstámogatás alapját képezi. Ugyanakkor számos pszichológiai kísérlet kimutatta, hogy az emberi bizonytalanságrevízió nem feleltethető meg teljes egészében e normatív bayesi módszernek. E jelenség pszichológiai elemzésére olyan vizsgálatok szolgáltak, melyekben objektíven számítható valószínűségi mértékeket vetettek össze ember által becsült, azaz szubjektív valószínűségekkel.

A vizsgálatok során mindhárom lehetséges emberi reakciót megfigyelték az új információkra:

(1) kezdeti bizonytalanságán gyorsabban változtat, mintsem a matematikai számítások szerint megengedhető lenne,

(2) „konzervatív módon" leragad a kezdeti bizonytalanságánál, és a kelleténél lassabban változtat,

(3) a bayesi számításokkal nagyjából egybevágóan folyik a bizonytalanság revíziója.

A valószínűségi ítéletalkotás vizsgálatakor az esetek többségében a (2) működési módot figyelték meg, s ezt a jelenséget konzervativizmus jelenségnek nevezték el. Különböző vizsgálatok azt mutatták, hogy az ember nem tudja megfelelően kihasználni a kezdeti bizonytalanságának megszüntetése érdekében az információk által kínált lehetőségeket. A bayesi számítások alapján elvárhatónál lényegesen kisebb mértékben csökkent az ember bizonytalansága. Számos kísérlet eredménye ugyanakkor arra utal, hogy a valószínűség-becslések újraértékelésekor lényegében a Bayes-tétel segítségével kiszámított valószínűségeknek megfelelő tendencia mutatkozott (bár lényegesen kisebb mértékben). Fontos megjegyezni, hogy a bayesi modell azon előfeltevésre alapoz, hogy relatíve függetlenek az egymás után következő adatok, míg az emberek (a mindennapi tapasztalatuk alapján) hajlamosak arra, hogy függőként kezeljék azokat. A konzervativizmus jelenséget több, egymástól eltérő, sőt egymásnak ellentmondó modell, elmélet próbálta magyarázni, de egyikük sem vált meghatározóvá.

4.2. Statisztikák becslése

A pszichológiai bayesi megközelítés abból a feltételezésből indul ki, hogy a spontán valószínűségi ítéletalkotások valamiféle számítások eredményeként állnak elő. Ez azt jelenti, hogy az ember tudatosan vagy tudattalanul olyasfajta fejszámolást végez a kapott információkkal, amely a Bayes-tételnek megfelelő algoritmusokhoz hasonló műveletekkel „képez" becsült valószínűségértékeket.
Tversky és Kahneman elmélete szerint azonban többnyire nem ez a helyzet [3]. Az általuk végzett vizsgálatok alapján a folyamat lényege, hogy az emberi tudat a számításokat heurisztikák alkalmazásával helyettesíti. Heurisztika alatt a pszichológiai megközelítésben kognitív folyamatok ismert, sajátos működési mechanizmusait értjük. Mérnöki szempontból ezeket leegyszerűsítve olyan ügyes módszereknek nevezhetnénk, amelyek valamilyen egyszerűsítés alkalmazásával elég jó megoldást adnak egy relatíve összetett vagy nehezen kezelhető problémára. Az alábbiakban áttekintjük azokat a területeket, ahol a heurisztikák valamilyen formában szerepet játszanak a valószínűségi ítéletalkotás során.

4.2.1. Elemi események becslése

Egy szituáció bizonytalan voltát az ember gyakran nem ismeri fel. A bizonytalanság kiiktatása leginkább olyan szituációkban jelenik meg, ahol jellemzőek a szélsőségesen

nagy, illetve a szélsőségesen kis valószínűségi értékek. Sajátos értékelési tendencia jelentkezik a valószínűségi skála végpontjain elhelyezkedő értékek becslésénél: az egészen kicsi valószínűségeket általában felülértékelik, a nagy valószínűségeket pedig alulértékelik. Ez a jelenség a szubjektív valószínűségbecslés terén létrejövő *centrális tendenciaként* ismert.

4.2.2. Az eloszlás becslése

A spontán ítéletalkotási mechanizmusok nincsenek tekintettel a statisztikai minta eloszlásából fakadó tulajdonságokra, annak következményeire. Általánosságban az emberek nem veszik figyelembe a minta nagyságát, ezért nagyjából azonos eloszlásokat feltételeznek például 10-es, 100-as, és 1000-es mintanagyságokra. Tekintsünk példaként egy éves statisztikát a koraszülöttek arányára az összes újszülött gyermekhez viszonyítva. Magyarországon ez az arány igen magas, átlagosan 8-9% százalék körüli. Ha feltesszük azt a kérdést, hogy várhatóan hol lesz nagyobb az átlagtól való eltérés: egy kisvárosban, ahol évente 1000 újszülött jön a világra, vagy egy faluban, ahol évente 10, akkor az emberek többsége a várost nevezné meg, nem pedig a falut, ahogy az valójában helyes lenne. A populáció egészét tekintve ugyanis az 1000 fős minta jóval reprezentatívabb, mint a 10 fős minta. Tehát statisztikailag az 1000 fő alapján számolt átlag közelebb helyezkedik el a teljes népességre vetített átlaghoz, mint a 10 fő alapján számított. A problémát a globális és a lokális reprezentáció lényegének összekeveredése okozza. Egy teljes populációra megadott statisztika értelemszerűen annak egészére vonatkozik, és ebben a kontextusban globális tulajdonságnak tekinthető. Egy adott alpopuláción ez a mennyiség lokálisan jelentősen eltérhet a globális átlagtól. Ezzel szemben az emberi ítéletalkotásba egyfajta eltúlzott lokális reprezentáció épült be, azaz a kisebb mintáról ugyanakkora reprezentativitást feltételezünk, mint egy megfelelően nagy mintáról. Az idevágó kísérletek eredményei meggyőzően azt mutatják, hogy a kis minták reprezentativitását a vizsgálatokban részt vevő egyének túlbecsülték, és becsléseik folyamán sorozatosan követtek el egészen elemi valószínűségszámítási hibákat. A mintareprezentativitás túlbecsülésének jelenségét Tversky nevezte el a *kis számok törvényé*nek [1].

4.2.3. A variancia becslése

A variancia egyfajta értelmezésben a heterogenitás mértékének tekinthető. Az elvégzett kísérletek arra mutatnak rá, hogy a homogenizálás irányába ható tendencia mutatkozik a heterogenitás becslésekor, vagyis a kísérleti személyek az egyes rendhagyó események lehetőségét nagyfokban limitálták becsléseik során. Az emberek a gyakorlati életben tett megfigyeléseik révén szekvenciálisan megtanulják -– adott binomiális sorozatok kimenetele alapján – az események valószínűségét. Az ennek igazolására végzett kísérletekben ez a tanulás feltűnően pontosnak bizonyult. Ezzel szemben az emberek többsége nehézséggel küzd a valószínűségi eloszlások varianciájának megbecsülésében. A kis valószínűségek megtippelése egyáltalán nem könnyű feladat, mert nagyon nehezen jutnak eszünkbe összehasonlítások, és a centrális tendencia eredményeképpen ezért kerüljük a valószínűségskála pólusainak használatát.
Ugyanakkor megfigyelhető egy ellentétes irányú hatás is, amelyet az átlaghoz való regresszió elvetéseként jellemezhetünk. Ennek az alapja az, hogy két egymást követő

időpontban végzett csoportos megfigyelés között a csoportátlagtól lényegesen eltérő értékek spontán közeledhetnek a csoportátlaghoz, azonban ezt az ember legtöbbször külső hatásnak tulajdonítja. Tekintsünk példaként egy időmérő edzéssorozatot, ahol öt egymást követő alkalommal mérik egy rögzített táv megtételének idejét húsz embernél. Tegyük fel, hogy a harmadik edzésen születtek a csoportátlagot jelentősen meghaladó eredmények, azonban ugyanezen emberek eredményei a negyedik edzésen már nem voltak jelentősen különbözőek a csoportátlagtól. Mások ellenben a negyedik edzésen szerepeltek jobban az átlagnál, és az ötödiken rosszabbul. Ilyen esetben az ember hajlamos a negatív irányú változást olyan eredményromlásként értelmezni, amelyet külső tényezők idéztek elő, holott lehet, hogy csak arról van szó, hogy a korábbi kiugró értékhez képest ismét az átlagos teljesítményt sikerült elérni. Tehát ekkor az átlaghoz való spontán visszatérés lehetőségét vetjük el.

4.2.4. A függetlenségre vonatkozó ítéletek

Az egyén mentális fejlődése során a statisztikai függetlenség gondolata egy bizonyos életkor után alakul ki. Nyilvánvalóan véletlenszerű szituációban hajlamosak vagyunk azt hinni, hogy adott események, melyek nem jelentkeznek egy ideig, a jövőben valószínűbben fordulnak elő. Tehát egyfajta korrekciós tendencia érvényesülését várjuk a véletlenszerű, független események egymás utáni sorában [1]. Ezt a jelenséget hívják a szerencsejátékos tévedésének (gambler's fallacy). Ez például egy olyan egyszerű eseménynél is tetten érhető, mint a pénzfeldobás. Tegyük fel, hogy szabályos érmével játszott pénzfeldobás játékban (p(fej) = p(írás) = 0,5) egymás után ötször *fej* lesz az eredmény. Ekkor a szerencsejátékos tévedésének hatása miatt úgy tűnhet, hogy ezután sokkal valószínűbb, hogy *írás* fog következni, mert „így állna helyre a rend", vagyis így közelítene az elvárt eloszláshoz a mért gyakoriság. A valóság azonban az, hogy a megelőző öt forduló eredménye nem befolyásolja a jelenlegi forduló kimenetelét, tehát azonos valószínűséggel következhet *fej* vagy *írás*.
A bizonytalanság kiiktatására szolgáló tendenciából természetszerűleg következik az, hogy a véletlenszerűség kiiktatására is irányul. Ez azt jelenti, hogy hajlamosak vagyunk ott is oksági és korrelatív összefüggések vélelmezésére, ahol ilyenek nincsenek. Másképp fogalmazva hajlamosak vagyunk olyan esetekben is törvényszerű együttjárások meglétét feltételezni, ahol erre sem a megfigyelt együttjárások gyakorisága, sem a tényezők közötti logikai kapcsolat nem nyújt elegendő okot. Ezt nevezzük illuzórikus korrelációnak.

4.3. Heurisztikák

Amikor az ember bizonytalan helyzetekben alkot ítéleteket, korlátozott információfeldolgozási kapacitása miatt egyszerűsítő eljárásokat kell, hogy alkalmazzon. Ennek következtében a valószínűségi ítéletalkotás egyes elméletek szerint az emberi ítéletalkotás néhány spontán működési szabályára épül, és nem aritmetikai műveletek sorának végrehajtására. Ennek megfelelően tehát heurisztikus jellegű. A valószínűségi ítéletalkotás sajátos jellemzőit Kahneman és Tversky szerint három heurisztika alkalmazása magyarázza: (1) a reprezentativitás, (2) a hozzáférhetőség és (3) a rögzítés és igazítás [3].

4.3.1. Reprezentativitás

A reprezentativitáson alapuló heurisztikákat akkor alkalmazzuk, amikor valamely egyed, esemény, osztály (X) egy nagyobb osztályhoz (M) való tartozásának valószínűségét kell megbecsülnünk. Ennek megfelelően négy alesetet különböztethetünk meg:

- (1) Eloszlás – jellemző érték: M egy osztály, X_i pedig az egyik az osztályban definiált változók (jegyek) közül, melynek egy konkrét értéke x_i. Ekkor az M osztályt reprezentáló x_{rep} érték egy releváns változó (X_r) adott osztályban előforduló értékeinek (x_r) a jellemző értéke, például átlaga. Egy egyszerű példa erre a magas férfiak osztálya, ahol az osztályt reprezentáló érték az átlagos magasság.

- (2) Kategória – egyed: M egy osztály, X az osztály egy egyede. X akkor reprezentálja M-et, ha olyan jegyekkel bír, mint az osztályba tartozó egyedek.

- (3) Populáció – minta: M egy osztály, X egy részhalmaza, azaz $X \in M$. A reprezentativitás ezen aspektusának a mintavételezésnél van szerepe.

- (4) Kauzális rendszer – hatás : M egy kauzális rendszer, X egy lehetséges következmény. X reprezentálja M-et vagy azért, mert erősen asszociáltak, vagy mert X tényleges (vagy vélt) következménye M-nek.

Tekintsünk egy példát a (2) esetre, mivel a reprezentativitási heurisztika itt érzékeltethető szemléletesen. *Mi a valószínűsége annak, hogy egy budapesti utcán szembejövő ember foglalkozását tekintve bankár?* A valószínűség értékelésében a példa szerinti esetben meghatározó szerepet játszik, hogy a szembejövő járókelő milyen mértékben reprezentál egy bankban dolgozó vezető beosztású embert, mennyire felel meg bankárokról kialakult elképzelésünknek. Tehát hasonlóság alapján döntünk, s ennek eredménye számos esetben élesen eltér a tételes valószínűségszámítás alkalmazásán alapuló döntéstől. A reprezentativitási heurisztika gyakran együtt jár bizonyos valószínűségi összefüggésekkel szembeni érzéketlenséggel, például a következőkkel:

- események elsődleges valószínűsége
- mintanagyság
- véletlen helytelen értelmezése
- becslés pontossága és megbízhatósága

Az emberek gyakran mutatnak érzéketlenséget az események elsődleges valószínűségével szemben. Ez megmutatkozhat például abban, hogy figyelmen kívül hagyjuk az előfordulási gyakoriságon alapuló valószínűséget, amikor a reprezentativitás szerepet játszik. Amikor megbecsüljük egy szembejövő járókelő foglalkozásának valószínűségét, elsősorban azt kellene meggondolnunk, mekkora a valószínűsége annak Budapest egy adott részén, hogy a kérdéses foglalkozású emberrel találkozunk. Egy bankárral való találkozás valószínűsége a belvárosban vagy általában bankok közelében jóval nagyobb, mint egy peremkerület lakóövezetében. A minta nagysága iránti érzéketlenség és a véletlen helytelen értelmezése is hasonló mechanizmusokból vezethetők le. Az emberek

az adott mintát mindig a populáció reprezentánsának fogják fel, függetlenül a minta nagyságától; s a nagy számok törvényének reprezentálását várják el a kis elemszámú mintáktól is (kis számok törvénye). Ez alapozza meg a szerencsejátékos tévedése (gambler's fallacy) jelenségét is. (Tversky és Kahneman szerint, a konzervativizmus jelenségére is magyarázatot ad a minta nagysága iránti érzéketlenség [1].)
A becslés pontossága és megbízhatósága iránti érzéketlenség egyfelől azt jelenti, hogy rendszerint figyelmen kívül hagyjuk, hogy egy adott esetben a becslés várható pontossága mekkora, azaz mekkora lesz a becsült érték konfidenciaintervalluma. Másfelől mellőzük a rendelkezésre álló evidenciák megbízhatóságát, és ezáltal az arra alapozott becslés megbízhatóságát. Ehelyett a jövőbeni események becslésénél egy adott minta reprezentatív voltát használjuk fel. A becslések megbízhatóságára vonatkozó kvantitatív mutatók iránti igényt ily módon csökkentik a reprezentativitáson alapuló jóslások.
A reprezentativitáson alapuló ítéletek mindemellett az érvényesség igen erős illúzióját keltik. Ennek lényege, hogy minél nagyobb a hasonlóság, a redundancia az evidenciák között, annál megalapozottabbnak véljük a döntést. Tehát az egymástól függő, az egymáshoz képest redundáns és korrelatív inputok alapján kialakított döntésekben jobban megbízunk, mint azokban, amelyeket egymástól független, egymáshoz képest nem redundáns, nem korrelatív inputok alapján hoztunk. Statisztikailag a független inputokon alapuló döntés a megbízhatóbb, mégis többre becsülünk egy függő információegyüttest annak belső konzisztenciája miatt. Ezáltal szembekerülünk az ítéletalkotás normatív előírásaival, azonban a reprezentativitáson alapuló heurisztikák alkalmazásakor a bemenetek belső konzisztenciája jobban hasznosítható, mint a statisztikai erő.
Kahneman és Tversky szerint a reprezentativitáson alapuló ítéletekben ragadhatók meg a regresszió téves értelmezésének okai [3]. A középarányos irányba való regresszióként ismert általános jelenség következménye, hogy bizonyos folyamatok esetében oszcilláló részkimenetek előfordulása várható. Ez a jelenség nehezen kezelhető az emberek számára. Gyakran feltételezik a regresszió létezését ott, ahol lényegében nincs (lásd például a szerencsejátékos tévedése), de nem vesznek róla tudomást ott, ahol létezik. Ugyanakkor általában külső oksági magyarázatot találnak rá, még ha fel is ismerik a regresszió jelenlétét. A jelenség nehezen kezelhetőségét az okozza, hogy nem egyeztethető össze sem azzal a felfogással, hogy a jósolt eredmény maximálisan reprezentatív a bemenetre nézve, sem azzal, hogy az eredményváltozó értékeinek annyira kell szélsőségeseknek lenniük, mint a bemeneti változók értékeinek.

4.3.2. Hozzáférhetőség

Akkor alkalmazzuk ezt a heurisztikát, amikor annak alapján becsüljük meg valamely esemény valószínűségét, hogy mennyire könnyen vagyunk képesek mnemikusan (emlékezési technikákkal) felidézni a rá vonatkozó példákat. A hozzáférhetőség alapján kialakított ítéletek (hasonlóan a reprezentativitáson alapulókhoz), jelentős statisztikai megbízhatóságot mutatnak, mert a könnyebb felidézhetőség általában az esetek múltbeli nagyszámú előfordulásán alapul. Természetesen más tényezők is szerepet játszhatnak a mnemikus anyag „hozzáférhetőségében", ezért e heurisztika alkalmazása ugyancsak sajátos, a valószínűségszámítás eredményeivel összehangolhatatlan ítéleti outputok forrása lehet. Kutatók állásfoglalása szerint különböző előítélettípusok eredetét találjuk meg az itt előforduló torzításokban:

a) példák felidézhetőségén alapuló torzítás,

b) keresési rendszer hatékonysága szerinti torzítás,

c) elképzelhetőségből fakadó torzítás,

d) illuzórikus korrelációból fakadó torzítás.

A nagy érzelmi töltésű tapasztalatok, illetve a közvetlen átélésből és a másodkézből való értesülésből adódó információk felidézhetőségének összehasonlítása esetében a felidézhetőségből adódó valószínűség-becslésbeli értékeltolódást tapasztalhatjuk. Például ha valaki személyesen szenvedett el egy balesetet, a baleset bekövetkezésének lehetőségét sokkal valószínűbbnek fogja tartani, mint az, aki csupán hallott ilyenekről. A felidézhetőséghez kapcsolódó további faktor az esemény időbeli elhelyezkedése. Egy távolabbi esemény becslést befolyásoló hatása ugyanis kisebb, mint egy közelmúltban történt eseményé.
A keresési rendszer hatékonysága miatti torzítás vélhetően annak a következménye, hogy bizonyos problémákhoz kialakítunk egy optimális keresési módot. Minél jobban eltér egy adott helyzetben a becslés felállításához szükséges információkeresés ettől az optimális módtól, annál nagyobb lesz a torzítás. Tekintsük egy olyan példát, melyben adott betűket tartalmazó szavak gyakoriságát kell megbecsülni. Ha feltesszük azt a kérdést, hogy az „e" betűt az első vagy a harmadik pozícióban tartalmazó szavak gyakorisága a nagyobb, akkor nagyobb valószínűséggel az előbbi opciót fogjuk választani. Ennek az oka az, hogy a szavakat rendszerint a kezdőbetűjük szerint rendezzük, keressük. Az első opció ennek megfelel, míg a második ettől eltér, és így nehezebb felidézni ilyen szavakat. A centrális tendencia jelenségért is a keresési rendszer hatékonysága a felelős. Ha egy személy olyan instrukciót kap, hogy nevezzen meg például valamely 0,005 vagy 0,995 valószínűségű eseményt, igen nehezen talál rá példát. Maga a jelenség arra vezethető vissza, hogy igen valószínűtlennek tartjuk a valószínűségskálán ennyire poláris elhelyezkedésű eseményekkel való találkozást.
Az elképzelhetőségből (imaginációból) fakadó torzítással azokban az esetekben kell számolnunk, amelyekben olyan osztályok gyakoriságát kell megállapítanunk, melyekről nincsenek közvetlen emlékképeink, hanem csak – bizonyos szabályok szerint – elképzelhetjük őket. Néhány példát általánosítunk ezzel a módszerrel, és arra támaszkodva állapítjuk meg a gyakoriságot és a valószínűséget. Erre egy egyszerű példa 10 fős mintában 3, illetve 7 fős diszjunkt részcsoportok képzése. Bár matematikailag az említett részcsoportok számossága megegyezik, a 3 fős csoportok képzése jobban elképzelhető, így ezt véljük gyakoribbnak.
Illuzórikus korrelációból fakadó torzítás akkor lép fel, amikor két jelenség között erős tudati asszociáció áll fenn, azonban ez statisztikailag nem állja meg a helyét. Ilyen esetekben akkor is gyakori közös előfordulást feltételezünk, amikor ennek a minta valójában ellentmond.

4.3.3. Rögzítés és igazítás

E heurisztika lényege abban ragadható meg, hogy a feladat elvégzése során az ítéletalkotó kijelöl valamilyen kezdő értéket, azt veszi alapul és ahhoz igazítja hozzá a végső döntést. A kiindulópont származhat a probléma megfogalmazásának módjából, vagy

partikuláris számításokból. A döntés és a cselekvés kivitelezhetősége szempontjából a rögzítésen és igazításon alapuló döntési manőver nagy jelentőséggel bír, és felelős olyan jelenségekért is, mint az elégtelen igazítás és a konjunktív és diszjunktív események értékelésében jelentkező típusos tévedések.

Annak alapján észlelhetjük az elégtelen igazításra irányuló tendenciát, hogy az ítéleti érték nagyságát túlzottan befolyásolja a kezdő érték megadása. Valószínűségi eloszlások becslésekor jelentősen eltérő értéket adunk meg, ha kiindulási érték nélkül kell megbecsülnünk egy adott esemény valószínűségének 90%-os konfidenciaintervallumát, szemben azzal, ha adott egy medián érték, és ehhez képest kell becslést adnunk az intervallumra. Akkor is ezt a jelenséget figyelhetjük meg, ha valaki befejezetlen számításokra alapozza az értékelését, tehát adott számú elemi művelet alapján becsli egy komplexebb művelet eredményét. Ha adott egy komplex műveletleírás két eltérő formában: $1 \cdot 2 \cdot 3 \cdots 10$ és $10 \cdot 9 \cdot 8 \cdots 1$, akkor az utóbbit nagyobbnak becsüljük, mert az első pár értékhez képest igazítunk.

A konjunktív és diszjunktív események értékelésében jelentkező típusos tévedések szintén az igazítás és rögzítés alkalmazásából adódnak. Tekintsünk egy példát három eseménytípus (elemi, diszjunktív, konjunktív) egymással történő összehasonlítására Engländer et al. nyomán.

Legyen az első egy elemi esemény, ahol egy olyan zsákból, amely fele részben tartalmaz piros és fele részben fehér üveggolyókat, egyszeri húzással egy pirosat vesz ki valaki. A konjunktív esemény például a következő: egy zsákban található üveggolyóknak 90%-a piros, 10%-a fehér. Valaki hétszer egymás után piros golyót húz ki a zsákból úgy, hogy a kihúzott golyót mindig visszateszi. Diszjunktív eseményre példa a következő: egy zsákból, amelyikben az üveggolyók 10%-a piros és 90%-a fehér, hét egymás utáni húzással valaki legalább egy piros golyót húz ki. A kihúzott golyót természetesen mindig visszateszi. Egy kísérletben a kísérleti személyek azt a feladatot kapták, hogy kössenek fogadást a három eseményfajtára, melyiket mennyire tartják valószínűnek. Kiderült, hogy a kísérleti személyek legnagyobb része a konjunktív eseményre fogadott, kisebb része az elemi eseményre, és legkisebb része a diszjunktívre, holott a diszjunktív esemény bekövetkezését tekinthetjük matematikailag a legvalószínűbbnek (0,52), utána az elemi eseményt (0,5) és legkevésbé valószínűnek a konjunktívat (0,48) [7].

Ezt a jelenséget Kahneman és Tversky a rögzítésből következő hatással magyarázzák [3]. Álláspontjuk szerint egy elemi esemény megállapított valószínűsége szolgál természetes kiindulópontként a valószínűség értékeléséhez mind a konjunktív, mind a diszjunktív eseményeknél. A kezdőponttól való igazítás általában nem kielégítő, ezért a végső értékelés mindkét esetben túl közel marad az elemi események valószínűségéhez. Ez a jelenség nem egyszer érezteti hatását a mindennapi gyakorlatban. A konjunktív események valószínűségének túlértékelésével találkozhatunk például a projekttervek teljes megvalósíthatóságának túlértékelésekor. Ha adott egy több részfolyamatból álló projekt, akkor a részfolyamatok együttes sikeres megvalósításának a valószínűségét rendszerint felülbecsüljük. A diszjunktív ítéleteknél tapasztalható alulértékelés pedig például a rendszerek részmeghibásodási lehetőségének alábecslésében mutatkozik meg. Vagyis bár az egyes részhibák egyenként jellemzően kis valószínűségűek, annak valószínűsége, hogy valamilyen hiba fellép, már nem elhanyagolható.

4.4. Torzítások a kockázat észlelésében

4.4.1. Perspektívahatás

A kockázat idői síkban történő érzékelése, és az ezen alapuló döntéshozatal egy további terület, ahol eltérés mutatkozik a valószínűségszámítás szabályaitól. Az adott időpillanathoz közelebbi kockázatot nagyobbnak, a távolabbit kisebbnek érzékeljük. Ezt a jelenséget a kockázat perspektívahatásának nevezzük. A perspektívahatást több kísérlet keretében vizsgálták. Ezek egyike egy szerencsejáték szituáció során demonstrálja a jelenséget [7]. A kísérlet résztvevőivel egy négyfordulós szerencsejátékot játszattak. Minden egyes fordulóban háromszor dobtak fel egy pénzdarabot és egyszer egy játékkockát. A résztvevőknek valamennyi dobás előtt tippelniük kellett, hogy mi lesz a dobás kimenetele. A pénzdarab feldobása esetén a lehetséges opciók: fej vagy írás, a kockadobás esetén pedig a lehetséges 6 szám egyike. Az egyes fordulók a dobások sorrendjében tértek el egymástól. Az első fordulóban a kockával kellett dobniuk először, majd az érmével háromszor. A második fordulóban a kockadobás a második helyre került. A harmadik és a negyedik fordulóban pedig rendre a harmadik és a negyedik helyre. A kísérlet további feltételei: a kísérleti személy csakis akkor nyer meg egy-egy fordulót, ha az érmefeldobások és a kockadobás kimenetelét is helyesen jósolja meg. Ha egyszer is téved, a forduló lejátszását megszakítják. Ha nyer, a tét 48-szorosát kapja, ha veszít, akkor a tét a vizsgálatvezetőé. A játék elkezdése előtt a játékosoknak valamennyi fordulóra fogadniuk kellett, és azt is meg kellett mondaniuk, hogy melyik fordulóban mekkora téttel fognak játszani. Ez lehetőséget adott annak mérésére, hogy a fordulókat egymáshoz képest mennyire tartják esélyeseknek. Bár a négy forduló során a nyerés valószínűsége objektíve teljesen azonos volt, a kísérletben résztvevők jelentősen eltérőnek ítélték az egyes fordulók kockázatát. Az eredmények azt mutatták, hogy minél korábban következett a kockadobás, a sorozat legnagyobb rizikójú tagja, annál nagyobbnak érzékelték a résztvevők az összkockázat nagyságát. Ennek megfelelően minél távolabb volt időben a kiemelt rizikójú pont, annál kisebbnek ítélték az összkockázatot.
Az elvégzett kísérletek eredményei alátámasztották azt a hipotézist, mely szerint minél később következik be a legnagyobb kockázatot képviselő esemény, annál esélyesebbnek tűnik az eseménysorozat egésze. Ebből a hipotézisből az is következik, hogy a valójában egyenlő valószínűségű sorozatokat az emberek nem egyenlőkként kezelik. A hipotézis hátterében megtalálható az adaptív célszerűség hatásának feltételezése: a biológiai adaptáció szempontjából az a legésszerűbb, ha döntési sorozatok esetében azt a lehetőséget választjuk, amelyikben a lehető legkésőbb jelentkezik a legnagyobb kockázat. Alapvetően csak ez a stratégia segíti a túlélést. Az, hogy a lényegében egyenlő esélyű és hasznosságú lehetőségeket egymástól különbözőkként észleljük, lehetőséget ad arra, hogy dönteni tudjunk közöttük, és ezzel megőrizzük cselekvőképességünket.

4.4.2. Egyenletesség

A kockázat időbeli észlelésének egy további emberi sajátossága a kockázat egyenletességének preferálása. Az emberi gondolkodás adott típusú szituációkban végzett kockázatbecslésnél előnyben részesíti az egyenletes kockázateloszlású sorozatokat. Különböző kísérletek igazolták, hogy ha az ember egy szituációban azonos összkockázatú esemény-

sorozatok értékelése alapján végez kockázati becslést, akkor az egymástól jelentősen különböző kockázatú elemeket tartalmazókkal szemben az azonos vagy egymáshoz igen közel álló kockázatú elemeket tartalmazó sorozatot preferálja.
Engländer és munkatársai egy rulettszerű kísérlettel demonstrálták ezt a jelenséget [7]. Adott volt egy sorrendben számozott szerencsekerék, amely 16 részre volt osztva 1-től 16-ig terjedő számozással és 4 színre volt festve (piros, sárga, kék, zöld). A kísérlet résztvevői számra (jele: N, $p = 1/16$), színre (jele: C, $p = 1/4$) és a számok páros, illetve páratlan voltára (jele: P, $p = 1/2$) fogadhattak. A kísérlet folyamán ezen eseményekből összeállított eseménysorozatok összkockázatát kellett megbecsülniük a kísérleti személyeknek. Például $N|P|P$ ($p = 1/16 \cdot 1/2 \cdot 1/2 = 1/64$) vagy $C|C|C$ ($p = 1/4 \cdot 1/4 \cdot 1/4 = 1/64$). Az eseménysorozatok valószínűsége minden esetben $1/64$ volt, mégis jelentősen többen ítélték az utóbbi eseménysort a legkevésbé kockázatosnak, mivel ennél volt a kockázat a legegyenletesebb.
A kockázategyenletesség preferenciája egyfajta „kvázi optimalizációs" módszernek is tekinthető, ugyanis az emberek könnyebben tudják kezelni azokat a helyzeteket, melyek jellemzői viszonylag tartósak. Ennek feltételezhetően az az oka, hogy a környezeti adaptáció szempontjából előnyös egy tartós tényező, mivel lehetővé teszi a hozzá történő alkalmazkodást. A jelenség úgy is értelmezhető, hogy az ember az összkockázatok nagyságának megítélésekor, a független részesemények valószínűségeinek összegzése során a matematikai gondolkodásnak jobban megfelelő szorzások helyett összeadáshoz hasonló műveletet alkalmaz.

4.4.3. Arányosság

Az összkockázat nagyságának meghatározását két további, egymással ellentétes hatású tényező befolyásolja, a kockázat sűrűsége és tartóssága. A kockázat sűrűsége azt fejezi ki, hogy az összkockázat milyen mértékben sűrített egyetlen vagy igen kevés számú pontba. A kockázat tartóssága a kockázat fennállásának tartamát jelöli. E két tényező hatását az egyenletesség vizsgálatánál leírt kísérlet egy változatával vizsgálták. A kíséreti alanyoknak három eseménysorozatra kellett kockázatbecslést végezniük, azaz fogadniuk téttel. Az első sorozat 1 elemet (N), a második 2-t ($C|C$), a harmadik pedig 4-et ($P|P|P|P$) tartalmazott. A számított összkockázat mindhárom sorozat esetében azonos volt ($N = 1/16$, $C \cdot C = 1/4 \cdot 1/4 = 1/16$, $P \cdot P \cdot P \cdot P = 1/2 \cdot 1/2 \cdot 1/2 \cdot 1/2 = 1/16$). Az „érzékelt" kockázat az elsőnél a legnagyobb, de csak egyszeri, míg a második és harmadik esetében lényegesen kisebb, de többször kell elviselni. A kísérlet eredményei azt mutatták, hogy ilyen szituációkban inkább választják az alacsonyabb mértékű kockázat többszöri elviselését (alacsony kockázatsűrűség, relatíve magas tartósság), mint az egyszeri nagy kockázatot (magas kockázatsűrűség, alacsony tartósság). Ilyen módon érvényesülhet az arányosság elve. (A környezeti adaptációban is inkább ez a megoldás vált sikeressé.)

4.5. Funkcionális referenciák

Kísérletsorozatok eredményei azt mutatták, hogy mind a laikusok, mind a felsőfokú matematikai képzettséggel rendelkezők a kockázati perspektívahatásnak megfelelően tették meg a fogadásaikat (habár a számított kockázati esélyek egyenlők voltak). Ha

az ítélet helyességének feltétlen kritériuma a normatív matematikai előírások alapján számított értékekkel történő egybeesés, akkor valóban hibáztak a kísérletben résztvevők. Ezzel szemben, ha a biológiai adaptáció vagy a mindennapi gyakorlat szemszögéből vizsgáljuk meg a kérdést, akkor nem ez a konklúzió adódik. A korábban említett szerencsekerék kísérletben résztvevőknek el kellett dönteniük, hogy a lehetséges opciók közül melyiket mekkora téttel fogadják meg. Tisztán formai szempontból ugyanúgy hibáztak volna, ha más sorrendet állítottak volna fel (az összekockázat ugyanakkora volt mindhárom esetben, tehát bármilyen sorrendezés hibás, amely valamelyik esetet nagyobb kockázatúnak jelöli a többinél). Ha elfogadják azt, hogy a rendelkezésre álló opciók egyenlő értékűek, és elutasítják a döntést, a szituatív feltételek alapján a lehető legrosszabb megoldást választották volna (kizárták volna magukat a játékból, vagyis elestek volna a potenciális nyereménytől). A kísérletek lebonyolítása során világossá vált, hogy működött a matematikai ismeretekből, illetve a valószínűségszámítási készségből származó kontroll (különösen a szakértők megjegyzései utaltak erre), de amikor döntöttek, nem a matematika normatív szabályaiból fakadó kritériumnak tettek eleget. A kísérletek arra mutattak rá, hogy a valószínűségi ítéletalkotás egy kettős, alternatív kritériumrendszeren alapul. Az adott feladathelyzet által mozgósított kritériumhalmaz orientálja az ítélet kialakításának módját és szabályait. A két kritériumtípus vagy más néven funkcionális referencia [7] a következő:

- *igazságkritérium* – a feladat elsődlegesen egy helyzet, állapot, tény igazságának feltárása,
- *hatékonysági kritérium* – egy cselekvés előkészítésére irányul az ítéletalkotás.

A hatékonysági kritérium azon esetekben kerül előtérbe funkcionális referenciaként, amikor cselekvéskényszerbe kerül az ítéletalkotó ember. A következő választásos típusú helyzetekben (jellemzően akkor, amikor nincs elegendő idő) adódhat elő cselekvéskényszer:

- azonos értékű lehetőségek között kell választani,
- túlságosan komplex a probléma az egyén információfeldolgozó teljesítményéhez viszonyítva,
- saját jelentőségéhez mérten túl összetett a probléma,
- elégtelen a rendelkezése álló információ mennyisége.

A funkcionális referenciának megfelelően két feladathelyzetet lehet megkülönböztetni, az egyik a „hagyományos gondolkodás” (például egy matematikai probléma helyes megoldása), a másik a becslés (például szerencsejátékban fogadás). A becslés közelebb áll az érzékelési folyamatoknál feltárt mechanizmusokhoz, a kognitív tevékenység analóg munkamódjai dominálnak, a működési szabályokban meghatározó a kockázat perspektívahatása, a heurisztikák alkalmazása, a mennyiségek logaritmikus konverziója.
Neumann szerint a digitális működés (matematikai számítások) elvitathatatlan előnye az elméletileg végtelen precizitás, pontosság, de nagy hátránya a labilis megbízhatóság (a nagy számú logikai lépést elvégző műveletsorok között egyetlen hiba is gyakori és

súlyosan hibás végeredményhez vezethet. Ezzel szemben az analóg működés (heurisztikák használata) kétségtelenül durva és pontatlan, de meglehetősen megbízható, és igen ritkák a súlyos, abszurd hibák).
Meglehetősen nehéz elméleti probléma a heurisztikus ítéletalkotás elhelyezése az ítéletalkotó folyamatok komplett rendszerében. A heurisztikus elméletet a filozófia oldaláról érkezett kritikák az emberi irracionalitás kísérleti bizonyítására való törekvésként jellemzik.

4.6. A kauzalitás szerepe

Bizonytalan szituációkban az emberi ítéletalkotás folyamatában a problémák időbeli szerkezetének fontos szerepe van. Bár tisztán matematikai szempontból a valószínűségi döntéshozatal időbeli paraméterei közömbösek, mégis számos kutatási eredmény alapozza meg azt az állítást. Az attribúciós (oktulajdonítási) elmélet kidolgozóinak (Kelley és McArthur) az az álláspontja, hogy a bizonytalan helyzetekben történő ítéletalkotás során is érvényesül az a tendenciózus kényszer, hogy az ember kauzális sémában értelmezze az egymást követő eseményeket [11], [10]. Jelölje X a rendelkezésre álló adatot (mintát), Y pedig a vizsgált eseményt, melynek valószínűségét becsüljük. Ekkor az alábbi sémákat különböztethetjük meg [2]:

- $X \rightarrow Y$: X kauzális információval szolgál Y-ról
- $Y \rightarrow X$: X diagnosztikus információval szolgál Y-ról
- $Y \leftarrow Z \rightarrow X$: X indikáció (adott szituációban információval szolgál Y-ról
- $Y -/- X$: X véletlenszerű Y szempontjából, tehát nincs sem direkt, sem indirekt kauzális él.

Ezek közül a kauzális iránynak van pszichológiai túlsúlya, ami azt vonja maga után, hogy a korábbi eseményt oknak, a későbbit okozatnak tekinti az ember. Ez olyan következménnyel is járhat, hogy akkor is ok-okozati összefüggés benyomása alakul ki az emberekben, mikor csupán az események egymásutániságáról van szó. Spontán, hétköznapi szituációkban a következtetések mechanizmusa úgy működik, hogy az időben előre (korábbiról későbbire) irányuló következtetéseket, mint okról okozatra, az időben visszafelé irányulóakat mint okozatról okra történő következtetéseket hajtjuk végre. Az előbbieket kauzális, az utóbbiakat diagnosztikus következtetéseknek nevezzük.
A kauzális következtetési forma természetesebb az ember számára, a környezeti adaptáció folyamatában ősibb követelményeket elégíti ki a diagnosztikushoz viszonyítva. A sikeres adaptációhoz, illetve a túléléshez szükséges gyors cselekvéses válaszok kivitelezését a kauzális következtetési forma biztosította. A múltbeli szituációk megértése és értékelése csak később, a biológiai és társadalmi fejlődés magasabb szintjén vált igazán értékesíthetővé. A diagnosztikus következtetési forma megjelenése már azt a típusú gondolkodást vetíti előre, amely elvezet a tudományos megismerés módszereihez.
Az elmélet kidolgozói a következőket feltételezik a hétköznapi ítéletalkotásban:

1. kézenfekvőbbeknek tűnnek és könnyebben végrehajthatók a kauzális következtetések, ezért jobban megbíznak bennük az emberek,

2. az információkból alapvetően a várható, jövőbeli történésekre vonatkozó következtetési lehetőségeket hasznosítjuk, és elhanyagoljuk az információk múltra vonatkozó, diagnosztikus értékét,

3. lényegesen könnyebben tudunk kauzális modelleket készíteni a bekövetkezett események megmagyarázására, mint diagnosztikusokat.

Mindezek jelentős következménye tudásmérnöki szempontból, hogy ha rendelkezésre áll olyan információ, amely egy kauzális következtetési lánchoz illeszkedik, akkor ez lesz a domináns. Mindazon mintákat, illetve a további információkat, melyek ehhez a sémához nem illeszkednek, figyelmen kívül hagyjuk. A modellalkotás szempontjából mindennek jelentős következménye van, akár hétköznapi értelemben, ha az embernek a környező világ mechanizmusaira modelleket alkotó tevékenységét tekintjük, akár formális értelemben, ha egy adott tárgyterület modellezését nézzük. A modellezési lépéseket három fő vezérlő elv köré csoportosíthatjuk: predikció, magyarázat, modellrevízió. Az első két elv a kauzális következtetés irányába illeszkedik, míg az utóbbi a diagnosztikus következtetéshez kapcsolódik. A kauzális irány túlsúlya a modell–kimenet viszonylat értelmezésében jelentkezik. Predikció esetében az ember hajlamos azon kimenetek súlyát eltúlozni, melyek leginkább illeszkednek a modellhez, míg a magyarázat jellegű megközelítésnél a modellnek azon elemeit súlyozzuk felül, melyek révén a kimenet létrejöhetett. Mindkét esetben az adott modell–kimenet illeszkedésének jóságára kerül a hangsúly. Ezzel szemben modellrevízió esetén annak megállapítása a cél, hogy a modell mely paramétereit kell módosítani a jobb illeszkedés érdekében. Ez utóbbi elv a diagnosztikus következtetést igénylő volta miatt háttérbe szorul, emiatt a valószínűségi ítéletalkotás és a kapcsolódó modellállítás folyamán túlsúlyba kerül a predikciós és a magyarázat alapú megközelítés.

A revízió elvetése nem csak a kauzális sémákhoz kapcsolódóan jelentkezhet. A hiedelmek melletti kitartás még új, annak ellentmondó adat esetében sem feltétlenül változik meg. Tegyük fel, hogy adott két csoport, A és B, melyeknek ugyanazon eseménynek a valószínűségéről eltérő hiedelmeik vannak. Ha egy olyan új evidencia válik elérhetővé, mely szerint mindkét hiedelem elfogadható, akkor optimális esetben közeledhetnek a hiedelmek egymáshoz, feltéve, hogy az új adat meggyőző, konzisztens a korábbiakkal. Ha az evidencia nem konzisztens a korábbiakkal, akkor a két csoport vagy elfogadja egymás hiedelmeit valamilyen mértékben vagy nem változtat az álláspontján. A valóságban azonban a hiedelmek nem hogy közelednének egymáshoz, hanem polarizálódnak, távolodnak egymástól. Az általános tendencia az, hogy minden információt a saját hiedelmeink megerősítésére használunk. Még abban az esetben is, amikor egy korábbi evidenciának megkérdőjeleződik a hitelessége (vagy kiderül, hogy egyértelműen tévedés), akkor sem csökken annyit a „hitünk”, mint amekkorát nő pozitív megerősítés hatására. Ennek mértékét a tudatban elraktározott korábbi siker-, illetve kudarcélményeink befolyásolják.

Ehhez kapcsolódik egy további jelenség, melyet visszatekintési torzításnak neveztek el. Ennek lényege, hogy egy vizsgált esemény bekövetkezte után, a visszatekintés folyamán túlbecsüljük azt, amit előre látni lehetett. Tehát a bekövetkezett eseményt szinte elkerülhetetlennek véljük utólag, azt feltételezzük, hogy mindent látni lehetett már előre, holott valójában nem ez volt a helyzet. Ennek hátránya az, hogy a későbbi predikciókat

túlzottan magabiztossá teheti, másfelől „eltakarja" az időközben fellépő hibák és más váratlan események jelentőségét.

4.7. A valószínűségi ítéletalkotás mint összetett szabályozó rendszer

Az emberi problémamegoldó gondolkodás egyik jellemző formája a valószínűségi ítéletalkotás folyamatában a heurisztikák alkalmazása, ami egyúttal rávilágít a következtetési folyamataink természetére is. A kutatók egy része azt feltételezi, hogy mindez csak a pszichológiai laboratóriumban végzett vizsgálatok melléktermeke, míg más jelentős csoportjuk ezt nem fogadja el. A valószínűségbecslések felülvizsgálata során az ember bayesi gondolkodásmódját hangsúlyozó felfogás ellentmondásban van a heurisztikus megoldások analóg, kvalitatív jellegével. A két neves kutató, Tversky és Kahneman a reprezentativitás heurisztikára vezeti vissza a konzervativizmus jelenséget [3]. Ők szakítanak azzal az előfeltevéssel, hogy az emberek „bayesi" mechanizmussal, azaz előzetes és feltételes valószínűségeket meghatározó adatok alapján tesznek kissé torzított, nem kellő hatékonyságú következtetéseket. E szerzők úgy vélekednek, hogy a becslések mechanizmusai az észleléshez közelebb álló szinten zajlanak, alapvetően a rendelkezésre álló információk, illetve az alternatív hipotézisek hasonlóságára vonatkozó komplex benyomásokon alapulnak. A konzervativizmus jelenséget abból vezetik le, hogy az emberek érzéketlenek egy lényeges statisztikai jellemzővel szemben, ez pedig a minta nagysága. A „bayesi módon gondolkodó ember" modelljének ugyancsak ellentmond a rögzítés és igazítás heurisztika rendszeres fellépésével kapcsolatos elvárás. Hasonlóan ellentmondásban áll egymással a valószínűség-becslések revíziójának hagyományos modellje és a műveletsorok kiszámításának elkerülésére irányuló törekvés. Bár meggyőzőnek tűnik a valószínűségi ítéletalkotás igen erős heurisztikus tendenciáinak létezése, a hagyományos bayesi kísérletek eredményei rámutatnak, hogy az egyének az információk növekedését valamilyen szintű hatékonysággal mégis csak hasznosítják. Az új információk – azaz a mintanagyság növekedése – befolyásolják a konzervativizmus mértékét a minta reprezentativitásától függetlenül is.

A valóságban a gondolkodás egyes folyamatai digitális lépéseinek megbízhatóságát jól értelmezik az analóg műveleteket alkalmazó heurisztikus becslések; a heurisztikák elnagyoltságát, logikai pontatlanságát pedig a digitális elemzések szükség esetén felismerik és korrigálják. Az ítéletalkotásunk során nem egymástól független, befejezett ítéleteket alakítunk ki. A különböző ítéletek között kölcsönös kapcsolatok állnak fenn, kontrollálják és kiegészítik egymást, több nézetből és különböző módokon közelítik meg ugyanazt a problémát. Az ítéletek összetett dinamikával működő hálózatokat, hierarchiákat alkotnak, melyekben ütköznek a különböző típusú heurisztikák és a normatív, digitális gondolkodási folyamatok eredményei. A valószínűségi megismerés intuitív és normatív mozzanatai egymást kölcsönösen kontrolláló és kisegítő műveletek. Tömören megfogalmazva, a következtetési, ítéletalkotási rendszer egy bonyolult visszacsatolási működést biztosító szervomechanizmust alkalmaz. A kiemelt kérdés az, hogy a következtetések intuitív és normatív elemei a valószínűségbecslések folyamatában hogyan hatnak egymásra.

A szakirodalom nagymértékben megosztott az emberi ítéletalkotás mint átfogó rend-

szer tekintetében. Az egyik irányzat szerint a nagy kockázatú helyzetekben az emberi döntések helyessége gyakran a túlélés alapfeltétele, ugyanakkor az ellenőrzésre és korrekcióra alig van vagy egyáltalán nincs lehetőség. Ebből az következne, hogy a legveszélyesebb szituációkban ki vagyunk szolgáltatva a tévedéseinknek, melyek a heurisztikák biasképző tendenciája miatt nem ritkák. Továbbá a különböző érzékelési modalitásokból származó információk visszacsatolási hatásaira, illetve a tanulási folyamatokra nem számíthatunk, mert túlságosan ritkák az ilyen rendkívüli helyzetek a helyes döntések és a megfelelő cselekvés elsajátításához. A szituációk kockázatos mivolta feloldhatatlan az ember számára, és a kockázatpercepciós kutatások szemlélete szerint az emberi pszichikum nincs felkészülve a kockázat kezelésére a visszacsatolás hiányában.

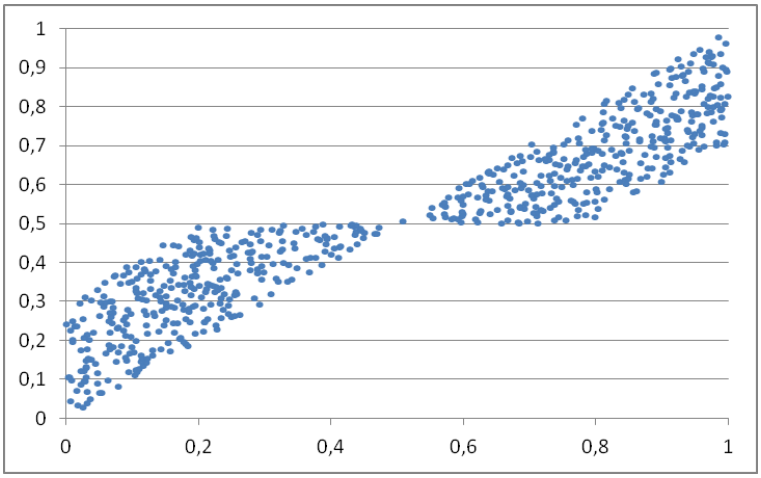
Egy másik irányzat szerint, azonban többek között biológiai megfontolások is indokolják a kételkedést a fentiekben. A leírt kísérletek alapján feltételezhető egy belső kontrollrendszer, amely az ítéletek megbízhatóságát biztosítja [7]. A különböző következtetési mechanizmusok szimultán működése, interaktív dinamizmusai, az analóg és digitális folyamatok, eltérő módon működő heurisztikák kölcsönösen ellenőrzik, ellensúlyozzák egymást. A valószínűségi problémamegoldás egységes rendszere az eltérő mechanizmusok nagyfokú együttműködéséből és a műveletek megfelelő hierarchikus szerveződéséből származik.

A fentiekkel ellentétben az extrém stressz kutatások eredményei arra utalnak, hogy létezik visszacsatolás a környezet és az egyén között. Az extrém kockázatú szituációkban azonnali reflexes reakciók indulnak be a veszélyeztető tényezők kiküszöbölésére. Ezzel párhuzamosan a kognitív apparátus értékeli a visszacsatolási információkat, és módja van a kockázatbecslési mechanizmusok alkalmazására (minél pontosabb képet rajzolnak az információk, és minél több idő áll rendelkezésre az ítéletalkotáshoz, a bizonytalanságrevízió során annál inkább tolódik át a gondolkodás a heurisztikák alkalmazásáról más, pontosabb kockázatbecslési mechanizmusokra). Továbbá szélsőséges szituációkban is elraktározódhat tapasztalat, főleg az erős érzelmi kontextussal rendelkező negatív tapasztalások esetén. Ez utóbbiak vésődnek be a legerősebben (persze ehhez túl kell élni a tévedést). Természetesen a más tapasztalatából való okulás mint információ nagyon fontos, sőt mint attitűd, alapja lehet a fejezet elején taglalt kockázatbecslési torzításoknak akkor, ha valóban nem vesszük figyelembe a mintanagyságot. A számunkra fontos emberek, csoportok tapasztalata nagyobb befolyással hat adott esetben, mint egy jóval nagyobb populáció kockázatbecslési ítélete.

A kockázatbecslés folyamatában nagy erővel hatnak a korábbi negatív tapasztalatok (sikertelen vagy kevéssé sikeres ítéletek következményei). Hasonló szituációkban befolyásolja, torzíthatja a kockázatbecslésünket még akkor is, ha az adott szituáció számos paramétere más, mint a korábbié (minden olyan helyzetben nagyobbnak fogjuk érzékelni a kockázat nagyságát, amely nagy vonalakban hasonlít egy korábbi, kudarcot eredményezett valószínűségi ítéletalkotásunkhoz). Ugyanakkor a rulettkísérletek során felismert torzítások is megjelennek a mindennapi gyakorlatban: ha tudottan 50-50%-os valószínűségi döntési helyzetben többször egyféleképpen döntünk sikertelenül, ragaszkodhatunk hozzá azon feltevés miatt, hogy az n-edik próbálkozásnál már nagyobb a valószínűsége annak, hogy igazunk lesz.

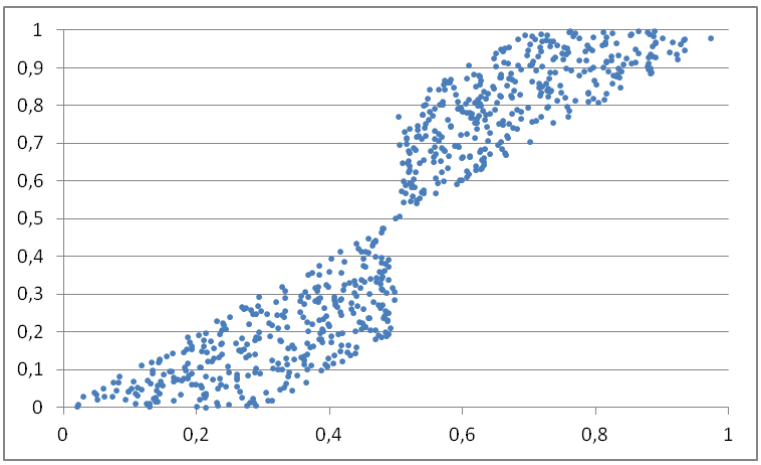
4.8. A torzítások hatása és azok kezelése

A valószínűségi következtetés és modellezés terén a paraméterek meghatározását illetően alapvetően kétféle következménye lehet az eddigiek folyamán ismertetett jelenségeknek, torzításoknak: túlzott magabiztosság (overconfidence) vagy a magabiztosság hiánya (underconfidence). A túlzott magabiztosság (lásd 4.1. ábra) a valószínűségi skála szélsőségei felé való eltolódást jelenti (azaz 0,5 alatt 0 felé, 0,5 felett 1 felé).



4.1. ábra. A túlzott magabiztosság (overconfindence) hatása valószínűségi következtető rendszereknél. A vízszintes tengely a túlzott magabiztosságot mutató becslést, a függőleges tengely az adat alapján számolt relatív frekvenciákat jelöli

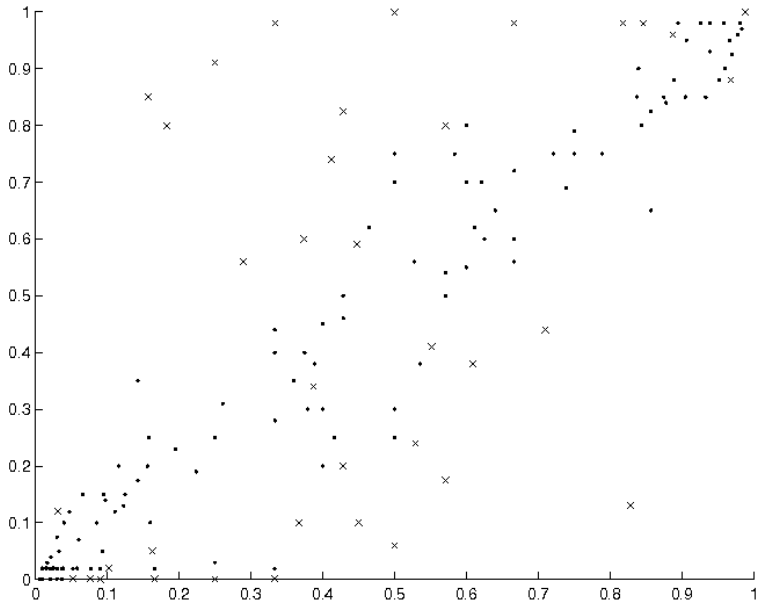
A magabiztosság hiánya (lásd 4.2. ábra) az előbbivel ellentétes tendenciát, azaz mindkét irányban $0,5$ felé való eltolódást mutat.



4.2. ábra. A magabiztosság hiányának (underconfindence) hatása valószínűségi következtető rendszereknél. A vízszintes tengely a magabiztosság hiányát mutató becslést, a függőleges tengely az adat alapján számolt relatív frekvenciákat jelöli

A heurisztikák miatt fellépő torzítások következtében a túlzott magabiztosság jellemző, amit célszerű figyelembe venni a valószínűségi modellezés folyamán. Ha a túlzott magabiztosságot az ideális becsléshez viszonyított szórásával (σ) jellemezzük, akkor

megállapítható, hogy kismértékű szórás estén ($\sigma < 0{,}2$) hatása a modell teljesítményére még minimális, azonban ezen érték felett jelentős romlás tapasztalható [8]. A magabiztosság hiánya ennél nagyobb romlást eredményez, azonban nem ez a jellemző a szakértői becslésekre. Onisko és Druzdzel vizsgálták a becslések pontatlanságának (kerekítésének) hatását egy valószínűségi következtető rendszerben [9]. Ennek során kimutatták, hogy a becsült valószínűségek 0-ra való kerekítése okoz nagymértékű teljesítményromlást.
A túlzott magabiztosság kezelése „kalibráció" segítségével lehetséges [5]. Ennek folyamán az szakértőt felkérjük adott valószínűségű események megbecslésére, majd a rendelkezésre álló adat alapján kiszámított relatív frekvenciákkal ezeket összevetjük. Ily módon átfogó képet kaphatunk arról, hogy az adott szakértő becslése mennyire szór az adatból számított értékekhez képest (lásd 4.3. ábra).



4.3. ábra. Kalibráció. A vízszintes tengely a szakértő becsléseit, a függőleges tengely az adat alapján számolt relatív frekvenciákat jelöli. A ponttal jelölt szakértői becslések az adat alapján számított valószínűség konfidenciaintervallumán belül találhatók, az x-szel jelölt becslések pedig a konfidenciaitervallumon kívül

Tudásmérnöki szempontból egy további következménye az ismertetett torzításoknak és heurisztikáknak, hogy a szakértő rendszerint elfogult lesz az adott problémához kapcsolódó konkrét információk súlyozásánál [4]. Ez azt jelenti, hogy ha rendelkezésre áll a problémához kapcsolódó általános információ és egyedi információ, akkor az utóbbi kap nagy hangsúlyt, míg az előbbit rendszerint elhanyagolja. Tekintsünk példaként egy építési projektet. Mivel számos hasonló vállalkozás lezajlott már, elérhető nagymennyiségű információ az ilyen típusú projektek eredményességéről, kritikus szakaszairól, a tervezett és a valós kivitelezési időről. Másképpen szólva eloszlás jellegű információhalmaz áll rendelkezésre. Ezzel állnak szemben a konkrét építési projekt egyedi információi, mint például a kivitelező vagy a helyszín paraméterei. A szakértő hajlamos a projekt tervezése folyamán az intuitív becslések meghozatalakor figyelmen kívül hagyni az eloszlásból származó információkat, és felülértékelni az egyedi informá-

ciókat. Ennek oka többrétű lehet, többek közt szerepet játszik az egyedi információk időbeli közelsége, a szakértő személyes érzelmi érintettsége a korábbi kudarcok vagy sikerek kapcsán, valamint a jellemző túlzott magabiztosság.
Elvileg lehetséges egy korrekciós metódust kialakítani az intuitív becslések javítására. Ennek lépései: (1) referenciaosztály meghatározása, (2) referenciaosztály paraméterei eloszlásának meghatározása, (3) intuitív becslés, (4) jósolhatóság becslése, (5) intuitív becslés korrekciója a jósolhatóság figyelembevételével. A gyakorlatban azonban ez többnyire nem megvalósítható, hiszen a legtöbb esetben egyfelől a referenciaosztály (amihez hasonlítjuk az adott problémát), másfelől a jósolhatóság (mennyire lehet pontos a becslés) meghatározása akadályokba ütközik. Emiatt inkább a torzítások lehetséges jelenlétének tudatosítása, és ebből kifolyólag az adott feladat több szempontú megközelítése nyújthat segítséget.

4.9. Összegzés

A matematikai szabályszerűségek, szigorúan meghatározott számítási elvek és módszerek szemszögéből tekintve, az emberi gondolkodásban, ítéletalkotásban, kockázatbecslésben, különösen a bizonytalan helyzetekben megmutatkozó, gyakran és szabályszerűen jelentkező elégtelenségek, torzulások, hibák ellenére a valószínűségi ítéletalkotás kognitív mechanizmusai a környezeti alkalmazkodás folyamataiban kellő hatékonysággal működnek. Olyan szituációkban, amelyekben precíz valószínűségi számítások elvégzésére az időkeret szűkössége nem ad lehetőséget, a becsléses ítéletalkotás sok esetben elnagyoltabb, durvább módszerei jobban szolgálják a biológiai túlélést. Az ember ilyen típusú kognitív működéseinek hét, az alkalmazkodást segítő különböző funkciója figyelhető meg (melyek matematikai nézőpontból gyakran minősülnek tévedésnek vagy hibának):
1. A cselekvőképesség biztosítása lerövidíti a döntéselőkészítés időtartamát azáltal, hogy strukturálja, értelmezhetővé teszi a másképpen nem kezelhető bemeneti adatokat, azaz megkülönböztethetővé tesz matematikai oldalról nézve egyenértékű opciókat.
2. Olyan műveleti eljárásokat alkalmaz (mint például a kockázati perspektíva hatás), amelyek az emberi becslésben azon kimeneteleket érzékeltetik esélyesebbnek (a túlélés szempontjából), melyekben a legnagyobb veszélyt hordozó események bekövetkezése a legkésőbbre várható.
3. Az emberi cselekvés természetes jellemzőihez adaptált következtetéseket, prognózist képez, vagyis a tartósan konstans paraméterekkel rendelkező szituációkat veszélytelenebbnek minősíti, mint a változó paraméterekkel bírókat.
4. Az emberi kognícióban egyaránt szerepelnek nagy pontosságú digitális és nagy megbízhatóságú analóg műveletek, ami lehetővé teszi a szituáció realisztikus szemléletét. Az analóg kognitív műveletek során az aktuális feltételek és körülmények összevetése a korábbi tapasztalatokkal (az ítéletalkotás általuk befolyásolt egyedi mechanizmusai) tereli az ítéletalkotást. A digitális műveletek során a kognitív mechanizmusok is az alapvető aritmetikai műveletekhez hasonlóan működnek.
5. Az emberi gondolkodás jellemzői kreatív megoldásokat tesznek lehetővé, melyek normatív következtetések révén nem, csupán új összefüggések felismerésével és alkalmazásával vezethetnek eredményre. Például: adott helyzetben a rendelkezésre álló adatokból pusztán hagyományos matematikai számításokkal nem található megoldás,

ugyanakkor más nézőpontból szemlélve az adathalmazt új kapcsolatok és összefüggések felismerésével (esetleg már korábban sikeresen megoldott feladatokban alkalmazott módszer adaptálásával) megoldhatóvá válik a feladat.
6. A diagnosztikus következtetésekben mutatkozó konzervativizmus egyfajta kiegyensúlyozó hatással bír a bizonytalanságrevízió során, szemben az ok-okozati következtetések lehatárolt mechanizmusaival.
7. A kockázatnagyság értékelésében, a becsléses és a kvázi matematikai gondolkodási folyamatok közötti párhuzamos együttműködés és kontrolltevékenység biztosítása az egyik legfontosabb kognitív funkció. A mechanizmus működésének lényege hasonlatos ahhoz a biológiai folyamathoz, amikor egy komplex érzékelési folyamatban az egyes érzékszervek által begyűjtött információk egymást pontosítva, korrigálva eléggé megbízható összképet alakítanak ki (például egy közeledő tárgy, jelenség veszélyességének megítélése során). Az emberi kockázatbecslések során nyilván előfordulnak tényleges hibák is, ezért az emberi ítéletalkotást, mint teljesítményt csak a fentiekben leírt folyamatok interaktív együttműködése teheti teljes értékűvé.

Irodalomjegyzék

[1] Tversky, A. and Kahneman, D. *Belief in the law of small numbers* in Kahneman, D. and Slovic, P. and Tversky, A., editors: *Judgment under Uncertainty: Heuristics and Biases*, pages 23-31. Cambridge University Press, New York, NY, 1982.

[2] A. Tversky and D. Kahneman. *Causal schemas in judgements under uncertainty.* In D. Kahneman, P. Slovic, and A. Tversky, editors, *Judgment under Uncertainty: Heuristics and Biases*, pages 117-128. Cambridge University Press, New York, NY, 1982.

[3] D Kahneman and A. Tversky. *Subjective probability: A judgement of representativeness, 1982.*

[4] D Kahneman and A. Tversky. *Intuitive prediction: Biases and corrective procedures.* In D. Kahneman, P. Slovic, and A. Tversky, editors, Judgment under Uncertainty: *Heuristics and Biases*, pages 414-421. Cambridge University Press, New York, NY, 1982.

[5] S. Lichtenstein, B. Fischhoff, and L. D. Phillips. *Calibration of probabilities.* In D. Kahneman, P. Slovic, and A. Tversky, editors, *Judgment under Uncertainty: Heuristics and Biases*, pages 306-334. Cambridge University Press, New York, NY, 1982.

[6] S. Lichtenstein, B. Fischhoff, and L. D. Phillips. *Calibration of probabilities.* In D. Kahneman, P. Slovic, and A. Tversky, editors, *Judgment under Uncertainty: Heuristics and Biases*, pages 335-354. Cambridge University Press, New York, NY, 1982.

[7] M. J. Druzdzel and A. Onisko. *The impact of overconfidence bias on practical accuracy of bayesian network models: An empirical study.* In *Working Notes of the 2008 Bayesian Modelling Applications Workshop, Special Theme: How Biased Are Our Numbers? Part of the Annual Conference on Uncertainty in Artificial Intelligence (UAI-2008), 2008.*

[8] T. Englander. *Viaskodás a bizonytalannal: A valószínűségi ítéletalkotás egyes pszichológiai problémái.* Akadémiai kiadó, Budapest, 1999.

[9] A. Onisko and M. J. Druzdzel. *Impact of quality of bayesian network parameters on accuracy of medical diagnostic systems. In AIME'11 Workshop on Probabilistic Problem Solving in Biomedicine (ProBioMed-11), 2011.*

[10] Leslie A McArthur. *The how and what of why: Some determinants and consequences of causal attribution. Journal of Personality and Social Psychology, 22(2):171-193, 1972.*

[11] Harold H Kelley. *The processes of causal attribution. American Psychologist, 28(2):107-128, 1973.*

5. fejezet

Egzakt, optimalizációs és Monte-Carlo-következtetések VGM-ben

5.1. Prediktív következtetés Bayes-hálókban

Mint a korábbi fejezetekben láttuk, a Bayes-hálók elsődleges célja, hogy az általuk reprezentált valószínűségi változók együttes eloszlását leírja. A Bayes-hálókban alkalmazott legalapvetőbb művelet a *prediktív következtetés*, amely során egy vagy több úgynevezett *célváltozó* eloszlását vagy adott konfigurációjuk valószínűségét keressük adott feltételek mellett. Ezek a feltételek tipikusan az úgynevezett *bizonyíték-* vagy *evidencia* változókra vonatkozó ismeretek.
Ebben a fejezetben a diszkrét változókat tartalmazó Bayes-hálókban való következtetéssel foglalkozunk. Az itt felölelt anyag legnagyobb része a Russel és Norvig mesterséges intelligenciáról írott könyvének [3] 14.4 (Egzakt következtetés Bayes-hálókban) és 14.5 (Közelítő következtetés Bayes-hálókban) szakaszait fedi le.

A bizonyíték típusai Ha egyértelműen ismert a bizonyítékváltozó által felvett érték, akkor *biztos evidenciáról*, ha csak az eloszlása ismert, akkor *bizonytalan evidenciáról* beszélünk. Látható, hogy a biztos evidencia a bizonytalan evidencia speciális esete: ekkor a változó eloszlása $0, \ldots 0, 1, 0, \ldots 0$ alakú.
A két legáltalánosabb következtetési eset az alábbi:

Egy célváltozó marginálisa. A következtetés alapfeladata: keresett az egyetlen célváltozó eloszlása a bizonyíték-ismeretek mint feltétel mellett ($\mathbf{P}(X|\mathbf{E} = \mathbf{e})$). A fő oka annak, hogy ezt a következtetési esetet külön kiemeljük, az, hogy sok következtetési algoritmus (elsősorban az egzaktak) esetében az összetett esetek kezelésének az a legegyszerűbb és leghatékonyabb, ha az összetett esetet visszavezetjük a célváltozó marginálisára.

Több célváltozó együttes konfigurációja. A következtetés általános esete: az evidenciák ismeretében keresett a (tipikusan) több célváltozó együttes eloszlása. Ha egy adott algoritmussal nem adható közvetlenül hatékony megoldás erre az esetre, akkor alkalmazható a láncszabály: vesszük az első célváltozót és kiszámítjuk

ennek valószínűségét a bizonyítékok mellett; ezután a bizonyítékok halmazát kibővítjük a célváltozó értékével, és a következő célváltozó valószínűségét már az új bizonyíték mellett számítjuk. Ha végigértünk a teljes célváltozó-halmazon, a keresett valószínűséget az egyes elemekre kapott részeredmények szorzataként kapjuk.

A fenti alapesetek alkalmazásaként még számos következtetési típust kaphatunk. Ezek közül a legjelentősebbek a következők:

Következtetés érzékenysége (Sensitivity of inference) Ebben az esetben azt vizsgáljuk, hogy adott változók értékére vonatkozó ismereteink, hogyan befolyásolják a célváltozók eloszlását. Azaz azt vetjük össze, hogy a célváltozó(k)nak az evidenciák eredeti állapota szerinti feltételes eloszlása hogyan változik, ha a bizonyítékok körét rendre kibővítjük a vizsgált változó értékeivel. A vizsgálat természetesen kibővíthető több vizsgált változóra is.

Többletinformáció értéke (Value of information) Ha a háló csomópontjainak (egy halmazának) konfigurációihoz valamiféle hasznosságértéket rendelünk, vizsgálhatjuk, hogy egy (vagy több) adott változó értékének ismerete hogyan fogja megváltoztatni a várható hasznosságot. Az információ értéke a következő képlettel számolható:

$$VOI(X|E) = \sum_{x \in X} P(x|E) \times |E_{\{V \setminus E \setminus X\}}[U(V)] - E_{\{V \setminus E\}}[U(V)]|; \tag{5.1}$$

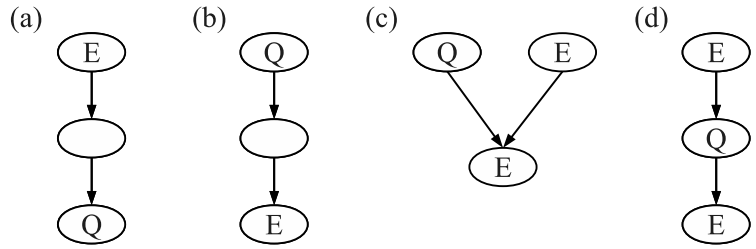
vagyis várhatóan (az új evidencia feltételes valószínűségével súlyozva) mekkora abszolút hasznosságváltozást fog okozni az új bizonyíték belépése.

5.2. A következtetési eljárások áttekintése

Bár a Bayes-hálókban való prediktív következtetés célja minden esetben alapvetően az, hogy bizonyos ismeretek birtokában a háló többi részére vonatkozó valószínűségi mennyiségeket számítsunk ki, magára a következtetés elvégzésére számos eljárás alkalmazható, aszerint, hogy mi a következtetés konkrét tárgya. Ennek megfelelően a következtetési eljárásokat többféle szempont szerint is csoportosíthatjuk, melyek közül a leglényegesebbek a következők:

- Az oksági viszonyok, azaz a különböző szerepű (megfigyelés-, illetve cél-) csomópontoknak a hálóban egymáshoz képest elfoglalt strukturális helyzete alapján,
- az alkalmazott eljárás jellege szerint,
- a Bayes-háló (és ezzel együtt a teljes következtetés) komplexitása szerint.

A megfigyelés- és a célcsomópontok oksági viszonya intuitíven fogalmazva a következtetés logikai irányát jelenti, vagyis azt, hogy a cél- és a bizonyítékváltozók között menő utakon milyen az élek irányítottsága. E szerint a besorolás szerint a fő csoportok a következők (az 5.1 ábra):



5.1. ábra. Tipikus következtetési esetek az evidenciák és célváltozók topologikus viszonya szerint: okozati (a), diagnosztikai (b), okok közti kimagyarázás (c) és kevert (d).

Okozati A megfigyelt csomópont a célpontnak az őse, azaz a következtetés iránya megegyezik a háló definiálása során feltett ok-okozati iránynak: az okot ismerve kell meghatároznunk az okozat eloszlását.

Diagnosztikai Az előző ellentettje, bizonyos következmények ismeretében kell meghatároznunk az azt kiváltó okokat (pontosabban azok eloszlását).

Okok közötti kimagyarázás A háló egy csomópontjának két őse (egy v-struktúra két felső csomópontja) között is fellép valószínűségi függés, ha a közös leszármazott értéke is ismert. Ilyen esetbe, azaz amikor egy másik ok ismerete a vizsgált ok eloszlását megváltoztatja, beszélhetünk egy ok kimagyarázásáról.

Kevert Általános esetben (amikor több megfigyelési és célpont változónk is van) a megfigyelések és célpontok strukturális viszonya nem fedhető le teljes mértékben egyik fenti kategóriával sem. Ilyenkor beszélünk *kevert* következtetésről.

5.2.1. A következtetési algoritmus

Az alábbiakban a legalapvetőbb, illetve a leggyakrabban alkalmazott következtetési eljárásokat soroljuk fel.

Egy teljes konfiguráció valószínűségének kiszámítása. Ha a megfigyelt - és a célváltozók halmazai együtt lefedik a háló egészét, a kérdéses valószínűség a Bayes-hálók definíciójául is szolgáló szorzatképlet egyszerű módosításával számítható:

$$P(\mathbf{V} = \mathbf{v}) = \prod_{X \in \mathbf{V}} P(X = x | Pa(X)), \tag{5.2}$$

azaz a változók egy topologikus sorrendjén végighaladva, az adott változónak rendre értékül adjuk a megfigyelt vagy lekérdezendő értéket. Ha a csomópont célváltozó, akkor annak feltételes valószínűségét bevesszük a szorzatba, ha pedig megfigyelt, kihagyjuk a szorzatból.
Bár ilyen következtetési eset a gyakorlatban szinte sohasem fordul elő, az eljárás megjelenik a következő következtetési módszer építőkockájaként.

Kimerítő felsorolás A legegyszerűbb következtetési eljárás: lényegében felsoroljuk a vizsgált bizonyíték-célváltozó konfigurációval kompatibilis teljes konfigurációkat és ezek valószínűségeinek összegeként adódik a keresett valószínűség. A részletes leírást az 5.3.1 alfejezet tartalmazza.

Következtetés fagráfokban Ha a vizsgált Bayes-háló struktúrája fagráf (vagyis bármely két csomópontja között maximum 1 út vezet), akkor lehetőség van a fentinél jelentősen hatékonyabb egzakt következtetésre. Ennek az algoritmusnak a leírását az 5.3.3 alfejezet tartalmazza.

Következtetés másodlagos struktúrában Ha a hálóban van irányítatlan kör, akkor a fagráfokra alkalmazható algoritmus nem használható. A másodlagos struktúrákat létrehozó algoritmusok ezt próbálják kiküszöbölni: a hálót úgy alakítják át, hogy az új modell az eredeti együttes eloszlást reprezentálja, de fastruktúrában, amelyben már hatékonyan következtethetünk. Természetesen ezeknél az algoritmusoknál a látszólagos komplexitás-csökkenésnek mindig megvan az ára, például az új struktúra csomópontjai az eredetihez képest jóval nagyobb értékkészletűek lesznek.
Az ebbe a kategóriába tartozó algoritmusok közül bemutatunk néhány egyszerűbbet (5.3.4 alfejezet), illetve egy összetettebbet, amely valós alkalmazásokban a leggyakrabban fordul elő (5.4 alfejezet).

Közelítő következtetés Monte-Carlo-eljárásokkal Ha a fenti egzakt eljárások nem alkalmazhatók (ennek oka a leggyakrabban a háló túl nagy volta), sztochasztikus szimuláció alkalmazható. Ezen eljárások alapötlete, hogy mintavételezzük a háló által reprezentált eloszlást, és a minta alapján számított statisztikával becsüljük a keresett mennyiséget (5.5.2 és 5.5.3 alfejezet).
Természetesen itt is igaz, hogy az egyszerűbb algoritmusok költsége a probléma növekedtével kezelhetetlenül nagyra nő. Az 5.6.4 fejezetben bemutatunk egy olyan Markov-lánc Monte-Carlo mintavételezési eljárást, amely nem közvetlenül a Bayes-háló együttes eloszlását mintavételezi (de belátható róla, hogy az egyensúlyi eloszlása ehhez tart), és így kezelhetőbb komplexitású megoldást kínál.

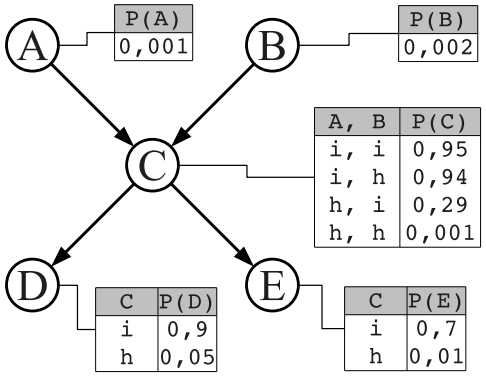
5.2.2. A következtetés komplexitása

Általános esetben (amikor a Bayes-háló tetszőlegesen nagy fokban összekötött részeket is tartalmazhat) belátható, hogy a következtetés feladata NP-nehéz, és a következtetés a csomópontok számában exponenciális komplexitású. Az ezzel kapcsolatos gondolatmeneteket és levezetéseket az 5.7 Függelék tartalmazza.

Következtetés polifákban. Fagráfokban (melyeknek bármely két csomópontja között maximum 1 út fut) belátható, hogy a következtetés végrehajtható polinom időben. Ez igaz mind az 5.3.3 alfejezetben bemutatott speciálisan fagráfokra alkalmazható algoritmusra (hiszen az eljárás minden egyes csomópontot maximum egyszer érint), mind az 5.3.2 alfejezetben bemutatott változó eliminációs algoritmusra. A részletes gondolatmeneteket a hivatkozott alfejezetekben közöljük.

5.3. Egyszerűbb egzakt következtető eljárások

A következőkben a Bayes-hálókban való egzakt következtetésre szolgáló alapvetőbb eljárásokat mutatunk be. Az algoritmusok többségét az alábbi modell felhasználásával illusztráljuk: A következő algoritmusok leírásánál minden esetben az egy változó mar-



5.2. ábra. A következtetési eljárások bemutatására használt egyszerű Bayes-háló.

ginális eloszlását kiszámító eljárást adjuk meg. Ha több változó együttes konfigurációjának valószínűségét akarjuk kiszámítani, az megtehető a *láncszabály* alkalmazásával: először kiszámítjuk az első célváltozó valószínűségét, ezután az evidenciák halmazát kibővítjük ezzel az értékkel és kiszámítjuk a második célváltozó valószínűségét. Ha ezzel az eljárással végigérünk az összes célváltozón, a menet közben kapott egyes valószínűségek szorzataként adódik a teljes konfiguráció valószínűsége.

$$P(X_1, \ldots X_n|E) = \prod_{i=1}^{n} P(X_i|E \cup \{X_j\}_{j=1}^{i-1}) \tag{5.3}$$

5.3.1. Következtetés felsorolással

A legegyszerűbb következtetési eljárásban lényegében felsoroljuk a szabad (nem cél- és nem evidencia) változók összes lehetséges konfigurációját, és azokat a célváltozó értékei szerint szeparálva összegezzük. Az egyes célváltozó-értékekhez tartozó összegek a célváltozó normalizálatlan eloszlását adják, amelyből a normalizált értékek már egyszerűen számíthatók:

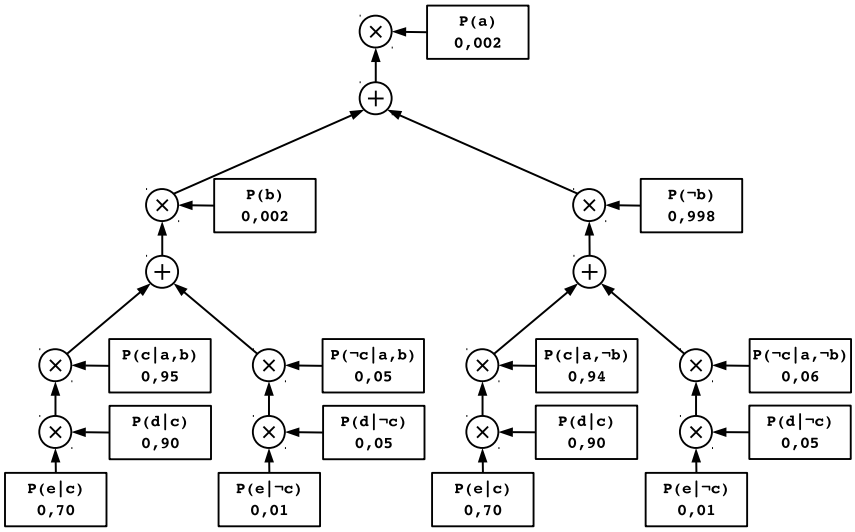
$$P(X = x|E = e) = \sum_{F_1} \cdots \sum_{F_n} \prod_{i=1}^{N} \underbrace{P(V_i|\{V_j\}_{j=1}^{i-1})}_{P(V_i|Pa(V_i))}, \tag{5.4}$$

ahol $\{F_i\}$ a „szabad”, azaz se nem cél-, se nem evidencia változók halmaza.
Valójában nem szükséges, hogy a szorzatban szereplő összes tag az összes szummán belül helyezkedjen el: ha a változók felett futó szummákat a Bayes-háló egy topologikus sorrendjében ($\prec_{top}$) helyezzük el, a $P(V_i|Pa(V_i))$ tagok előrehozhatók közvetlenül a vonatkozó $\sum_{V_i}$ szumma mögé, hiszen az értékük nem függ a szumma indexváltozójától,

így konstansként kiemelhetők. Így a kifejezés a következő lesz:

$$P(X=x|E=e) = \prod_{i:V_i \prec_{top} F_1} P(V_i|Pa(V_i)) \times \sum_{F_1} \Bigg(\prod_{i:V_i \prec_{top} F_2} P(V_i|Pa(V_i)) \times \sum_{F_2} \big(\ldots\big) \Bigg) \tag{5.5}$$

Magát a kifejezést és annak kiértékelését az 5.3 ábra szemlélteti, kiszámítása pedig az 5.1 algoritmus szerint történhet.



5.3. ábra. A felsorolásos következtetés által kiértékelt kifejezés szerkezete a $P(A|D = i, E = i)$ lekérdezés esetén.

A Felsorolás-Kérdezés csak annyit tesz, hogy rendre kibővíti az evidenciákat a célváltozó értékeivel, ezen inicializáló lépések után pedig Felsorol-Mindet számolja ki az 5.5 egyenletben szereplő összegeket, azaz a normalizálatlan eloszlás-vektor elemeit.

5.3.2. Következtetés változó eliminációval

Az előbbiekben bemutatott felsorolásos eljárás fő gyengesége hatékonysági szempontból, hogy a kifejezés-fa alsó csomópontjaiban található mennyiségeket feleslegesen, többször is kiértékeli. Természetesen adódik tehát a hatékonyságot javító módosítás: a felsorolásos módszerben többször is kiértékelt mennyiségeket tároljuk, és használjuk fel őket újra, amikor szükségesek lesznek.

A *változó eliminációs eljárás* által elvégzett feladat tehát a következő: adott egy Bayes-háló és ezzel együtt az általa ábrázolt $P(\mathcal{U})$ eloszlás; keressük a $P(X|E)$ eloszlást, ahol X a célváltozók, E az evidenciák halmaza. Az eljárás a számítást úgy hajtja végre, hogy $\mathcal{U}$-ból egymás után eliminálja a sem X-ben, sem E-ben nem szereplő változókat. Példaképpen tekintsük a $P(A|d,e)$ mennyiség kiszámítását az 5.2 ábrán bemutatott minta hálózatban. Számítsuk ki az alábbi kifejezést:

$$P(A|d,e) \propto \underbrace{P(A)}_{A} \sum_{b} \underbrace{P(b))}_{B} \sum_{c} \underbrace{P(c|A,B)}_{C} \underbrace{P(d|c)}_{D} \underbrace{P(e|c)}_{E}. \tag{5.6}$$

Algorithm 5.1 Egzakt következtetés Bayes-hálókban felsorolással

```
1: function FELSOROLÁS-KÉRDEZÉS(X, e, bn)                    ▷ returns P(X)
                                                  ▷ X : a lekérdezés változója
                                          ▷ e : az E változók megfigyelt értékei
                                                  ▷ bn : a valószínűségi háló
2:     Q(X) ← üres eloszlás
3:     for each x_i in X értékei do
4:         terjeszd ki e-t x_i-vel
5:         Q(x_i) ← FELSOROL-MINDET(VÁLTOZÓ[bn] e)
6:     end for each
7:     return NORMALIZÁL(Q(X))

8: function FELSOROL-MINDET(változók, e)                     ▷ returns valós szám
9:     if ÜRES(változók) then
10:        return 1.0
11:    Y ← ELSŐ(változók)
12:    if Y értéke y ∈ e then
13:        return P(y|Pa(Y))×FELSOROL-MINDET(MARADÉK(változók), e)
14:    else
15:        return Σ_y P(y|Pa(Y))×FELSOROL-MINDET(MARADÉK(változók), e ∪ {Y = y})
```

A kifejezésben minden tag-valószínűséget megjelöltünk a hozzá kapcsolódó változó nevével, ezek az úgynevezett *tényezők* (*factors*). A kiértékelés jobbról balra haladva, a következőképpen történik:

E: Mivel E változó értéke rögzített (a bizonyítékban) ezért itt nincs szükség E szerinti összegzésre, pusztán tároljuk E valószínűségét mint a feltételek (ebben az esetben C) függvényét egy vektorban: $f_E(E) = (P(e|c), P(e|\overline{c}))$.

D: Itt E-hez hasonlóan kételemű vektort kell tárolnunk: $f_D(D) = (P(d|c), P(d|\overline{c}))$.

C: Ebben az esetben a $P(c|A, b)$ értékeket egy $2 \times 2 \times 2$-es „mátrixban”, $f_C(C, A, B)$-ben tároljuk.

$\sum_c$: A C feletti összegzést két lépésben végezhetjük el: először összeszorozzuk a tényezőket, majd a szorzás eredményéből C feletti összegzéssel eliminájuk C-t. A tényezők szorzására az úgynevezett *pontonkénti szorzás* művelet szolgál (részleteit lásd lentebb). Az összegzés után az $f_{\not C DE}(A, B)$ tényezőt kapjuk, amely tehát D és E valószínűségeit tartalmazza, A és B szerint indexelve, C-t pedig összegezve (erre utal a $\not C$ alsó index).

... A további lépések már a fentiekből sejthetők: tároljuk f_B-t, majd összeszorozzuk az előbb kapott tényezővel, és B felett összegezve kapjuk $f_{\not B \not C DE} = \sum_b f_B \times f_{\not C DE}(A, B)$-t. Ezután már csak az A-hoz tartozó tényezővel kell megszorozni az eddig kapott eredményt: $P(A|d, e) \propto f_{\not A \not B \not C DE} = f_A \times f_{\not B \not C DE}$

Pontonkénti szorzás. A fenti számításokban használt tényezők (faktorok) valójában egy változóhalmaz konfigurációival indexelt táblázatok, illetve tömbök. A *pontonkénti szorzás* két ilyen táblázatot kombinál össze: az indexáláshoz használt változók halmaza a szorzatban a két eredetinek az uniója lesz, az egyes bejegyzések pedig az új indexből vetítéssel visszakapható eredeti indexekhez tartozó bejegyzések szorzata lesz, például $f_{1\cdot 2}(a,b,c) = f_1(a,b)f_2(b,c)$. Az 5.1 táblázat két tényező pontonkéni szorzatára ad példát.

A	B	$f_1(A,B)$	B	C	$f_2(B,C)$	A	B	C	$f_{1\times 2}(A,B,C)$
I	I	$0,3$	I	I	$0,2$	I	I	I	$0,3\cdot 0,2$
I	H	$0,7$	I	H	$0,8$	I	I	H	$0,3\cdot 0,8$
H	I	$0,9$	H	I	$0,6$	I	H	I	$0,7\cdot 0,6$
H	H	$0,1$	H	H	$0,4$	I	H	H	$0,7\cdot 0,4$
						H	I	I	$0,9\cdot 0,2$
						H	I	H	$0,9\cdot 0,8$
						H	H	I	$0,1\cdot 0,6$
						H	H	H	$0,1\cdot 0,4$

5.1. táblázat. Két tényező pontonkénti szorzata.

Az eljárás komplexitása polifákban.

A változó eliminációs algoritmusban egy-egy összegzés elvégzésével mindig eltávolítunk egy csomópontot a képletből, vagyis az ilyen összegzések számát a változók száma felülről korlátozza. Belátható, hogy a teljes eljárás költségét az határozza meg, hogy egy-egy összegzés során mekkora $f_{\ldots}(\ldots)$ tényezőt kell előállítani és feldolgozni. Az $f_{\ldots}(\ldots)$ tényezők mérete erősen függ az eliminálás sorrendjétől. A sorrend helyes megválasztása esetén, azaz ha az konzisztens a háló egy topologikus sorrendezésével, a legnagyobb ilyen tényező is arányos lesz az éppen eliminált csomópont CPT-jének méretével.
Ha tehát a hálóban a csomópontonkénti szülők számára, és a csomópontok értékkészletére is adott egy felső korlát (k és d), akkor az eljárás komplexitása $\mathcal{O}(nd^k)$ lesz (ahol n a csomópontok száma).

Változók relevanciája a következtetés szempontjából

Tekintsük a következő lekérdezést az 5.2 ábra hálójában:

$$P(D|a) \propto P(a)\sum_b P(b)\sum_c P(c|a,b)P(D|c)\sum_e P(e|c). \tag{5.7}$$

A $\sum_e P(e|c)$ mennyiség definíció szerint 1 lesz, vagyis belátható, hogy minden levélcsomópont, amely nem bizonyíték és nem célváltozó eltávolítható a hálóból a konkrét következtetési esetben, mivel nem befolyásolja annak kimenetelét. Természetesen, ha egy levélcsomópontot ily módon eltávolítottunk, az új háló is tovább „nyeshető”, amiből a következő tétel adódik.

5.1. tétel (Változók irrelevanciája)**.** Egy Bayes-hálóban minden csomópont, amely nem eleme vagy őse a bizonyíték- vagy a célváltozóknak, *irreleváns* az adott lekérdezés szempontjából.

A releváns csomópontok halmaza tovább szűkíthető a következők figyelembevételével:

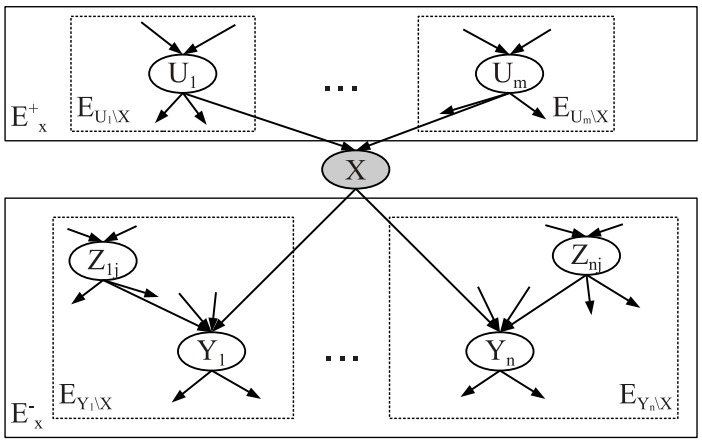
5.1. *definíció* (Morális gráf). Egy DAG (irányított körmentes gráf) *morális gráfja* úgy kapható meg, hogy az eredeti gráfban az egyes csomópontok szüleit egymással összekötjük, majd az így kapott gráfban töröljük az élek irányítását.

5.2. *definíció* (m-szeparáció). Az $\mathbf{X}$ és $\mathbf{Y}$ változóhalmazokat $\mathbf{Z}$ *m-szeparálja* a G DAG-ban, ha $\mathbf{Z}$ szeparálja $\mathbf{X}$-et és $\mathbf{Y}$-t G morális gráfjában.

5.2. tétel (Változók irrelevanciája II.)**.** Azon változók, amelyeket G-ben a bizonyítékok m-szeparálnak a célváltozóktól, az adott lekérdezés szempontjából *irrelevánsak*.

5.3.3. Következtetés polifákban

Ha a Bayes-háló egy úgynevezett erdő, azaz egy (potenciálisan) több fából álló gráf, akkor létezik a változók számában lineáris következtetési algoritmus. Ahhoz, hogy ezt belássuk, tételezzük fel, hogy a feladat egy célváltozó egy adott értéke valószínűségének kiszámítása.*



5.4. ábra. Következtetés polifában: az X célváltozó és az általa feltételesen függetlenné szeparált hálórészletek viszonya.

A feladat tehát a $P(X|E)$ valószínűség kiszámítása, amely a Bayes-szabály segítségével a következő alakra hozható, ha az evidenciák teljes halmazát kettébontjuk az X „feletti”

*Mind a teljes marginális eloszlás, mind egy többváltozós konfiguráció valószínűségének kiszámítása megoldható hasonló nagyságrendben mint amilyen egyetlen változó egyetlen értéke valószínűségének kiszámításához szükséges, hiszen a marginális eloszlás kiszámítható az eljárás egyszerű megismétlésével a változó összes értékére, egy többváltozós konfigurációra pedig alkalmazható a láncszabály: kiszámítjuk az első célváltozó valószínűségét, ezután felvesszük az evidenciák közé, és kiszámítjuk a második célváltozó valószínűségét az új, bővített evidencia mellett, majd végül a kapott valószínűségeket összeszorozva kapjuk az eredeti lekérdezésre a választ. Vagyis a következtetés nagyságrendje $\mathcal{O}(cn)$, illetve $\mathcal{O}(n^2)$ nagyságrendben lesz, ahol c az adott változó kardinalitása, n pedig a változók száma

(E_X^+) és X „alatti” (E_X^-) részekre:

$$P(X|E) = P(X|E_X^+, E_X^-) = \frac{P(E_X^-|X, E_X^+)P(X|E_X^+)}{P(E_X^-|E_X^+)}. \tag{5.8}$$

A $P(E_X^-|E_X^+)$ értéke kezelhető konstansként, így elhagyható, hiszen az X feletti eloszlást majd X értékei felett kapott értékek normalizálásaként kapjuk. Továbbá mivel X d-szeparálja E_X^+-t és E_X^--t, az egyenlet a következőképpen egyszerűsíthető:

$$P(X|E) \propto P(E_X^-|X)P(X|E_X^+). \tag{5.9}$$

$P(X|E_X^+)$ **kiszámítása** A $P(X|E_X^+)$ mennyiséget egyszerűen kiszámíthatnánk X szüleinek ($\mathbf{U}$) eloszlásai ismeretében, a $P(X|U)$ feltételes valószínűségek $P(U)$ szerint súlyozott összegeként. Ezt felhasználva $P(X|E_X^+)$ tovább bontható:

$$P(X|E_X^+) = \sum_u P(X|u, E_X^+)P(u|E_X^+). \tag{5.10}$$

Mivel X d-szeparálja $\mathbf{U}$ elemeit, és E_X^+ azokhoz tartozó részeit (az 5.4 ábrán E_X^+-nak az $E_{U_i \setminus X}$ dobozokban lévő részeit), $P(u|E_X^+)$ felírható a $P(u_i|E_X^+)$ mennyiségek szorzataként; a $P(u_i|E_X^+)$ valószínűségekben pedig E_X^+ a szűkebb $E_{U_i \setminus X}$ halmazokra cserélhető (ahol $E_{U_i \setminus X}$ az evidenciák azon halmazát jelöli, amelyek U_i-ből a Bayes-háló struktúra élein lépkedve elérhetők X érintése nélkül):

$$P(X|E_X^+) = \sum_u P(X|u) \prod_i P(u_i|E_{U_i \setminus X}). \tag{5.11}$$

Ez a képlet az eredetibe visszaírva már alkalmas lesz arra, hogy egy rekurzív eljárás alapját képezze, hiszen a $P(X|u)$-k a Bayes-háló CPT-inek bejegyzései, $P(u_i|E_{U_i \setminus X})$ pedig ($P(X|E)$)-hez hasonlóan számítható. $P(u_i|E_{U_i \setminus X})$ kevesebb evidenciát tartalmaz mint ($P(X|E)$), vagyis intuitíve belátható, hogy az algoritmus ezen része terminálni fog.

$P(E_X^-|X)$ **számítása** Kihasználva a háló struktúrájának fagráf voltát (hasonlóképpen az 5.11 egyenletre vezető megfontolásokhoz) $P(E_X^-|X)$ is felírható egy X gyermekei (Y_i) feletti szorzatként:

$$P(E_X^-|X) = \prod_i P(E_{Y_i \setminus X}|X). \tag{5.12}$$

Ha az előző képlet tagjait felírjuk X gyerekei (Y_i) és azok szülei ($\mathbf{Z}_i$) feletti összegzésekként, a következőket kapjuk:

$$P(E_X^-|X) = \prod_i \sum_{y_i} \sum_{\mathbf{z}_i} P(E_{Y_i \setminus X}|X, y_i, z_i)P(y_i, z_i|X). \tag{5.13}$$

Ha $E_{Y_i \setminus X}$-et E-hez hasonlóan felbontjuk Y_i alatti és feletti részekre ($E_{Y_i}^-$ és $E_{Y_i \setminus X}^+$), az alábbiak adódnak:

$$P(E_X^-|X) = \prod_i \sum_{y_i} \sum_{\mathbf{z}_i} P(E_{Y_i}^-|X, y_i, z_i)P(E_{Y_i \setminus X}^+|X, y_i, z_i)P(y_i, z_i|X). \tag{5.14}$$

Kihasználva, hogy Y_i d-szeparálja $E^-_{Y_i}$-t X-től és z_i-től, illetve, hogy z_i d-szeparálja $E^+_{Y_i\setminus X}$-t X-től és y_i-től, valamint kiemelve egy z_i-től nem függő tagot a második szummából, a következőkhöz jutunk:

$$P(E^-_X|X) = \prod_i \sum_{y_i} P(E^-_{Y_i}|y_i) \sum_{\mathbf{z}_i} P(E^+_{Y_i\setminus X}|z_i) P(y_i, z_i|X). \tag{5.15}$$

Alkalmazva a Bayes-szabályt $P(E^+_{Y_i\setminus X}|z_i)$-re:

$$P(E^-_X|X) = \prod_i \sum_{y_i} P(E^-_{Y_i}|y_i) \sum_{\mathbf{z}_i} \frac{P(z_i|E^+_{Y_i\setminus X}) P(E^+_{Y_i\setminus X})}{P(z_i)} P(y_i, z_i|X), \tag{5.16}$$

amiből $P(y_i, z_i|X)$-re alkalmazva a láncszabályt kapjuk, hogy:

$$P(E^-_X|X) = \prod_i \sum_{y_i} P(E^-_{Y_i}|y_i) \sum_{\mathbf{z}_i} \frac{P(z_i|E^+_{Y_i\setminus X}) P(E^+_{Y_i\setminus X})}{P(z_i)} P(y_i|X, z_i) P(z_i|X). \tag{5.17}$$

Itt kihasználhatjuk, hogy a d-szeparáció miatt $P(z_i|X) = P(z_i)$. $P(E^+_{Y_i\setminus X})$-t átnevezve β_i normalizációs állandóvá, a következőt kapjuk:

$$P(E^-_X|X) = \prod_i \sum_{y_i} P(E^-_{Y_i}|y_i) \sum_{\mathbf{z}_i} \beta_i P(z_i|E^+_{Y_i\setminus X}) P(y_i|X, z_i). \tag{5.18}$$

A β_i-k kiemelhetők az összegzések elé, az így keletkező $\prod_i \beta_i$ pedig elhagyható a későbbi normalizáció miatt. Y_i szülei (Z_{ij}-k) függetlenek lesznek egymástól Y_i ismeretében, így együttes valószínűségük szorzatként felírható:

$$P(E^-_X|X) \propto \prod_i \sum_{y_i} P(E^-_{Y_i}|y_i) \sum_{\mathbf{z}_i} P(y_i|X, z_i) \prod_j P(z_{ij}|E_{Z_{ij}\setminus Y_i}). \tag{5.19}$$

A fenti kifejezés már szintén felhasználható a rekurzív algoritmusban, hiszen

- $P(E^-_{Y_i}|y_i)$ a rekurzív megjelenése $P(E^-_X|X)$-nek;
- $P(y_i|X, z_i)$ CPT bejegyzések a hálóban;
- $P(z_{ij}|E_{Z_{ij}\setminus Y_i})$ pedig a rekurzív megjelenése $P(X|E)$-nek.

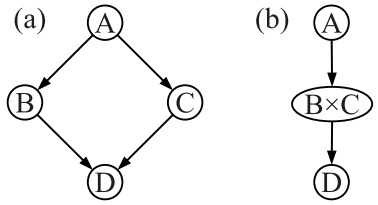
Az eljárás komplexitása. Látható, hogy az eljárás minden csomópontot maximum egyszer látogat meg, hiszen az a célcsomópontból indított rekurzív hívásokból áll, amelyek az élek mentén haladnak a célcsomóponttól távolodva (a rekurzív hívások során az azt indító előző csomópont kizáródik a meglátogatott csomópontok köréből); ebből pedig következik, hogy komplexitása a csomópontok számában lineáris.

5.3.4. Következtetés nem fa gráfokban

Ha a hálóban több út is vezet két csomópont között, az előző szakaszban tárgyalt algoritmus nem alkalmazható. Ilyen helyzetben alkalmazhatunk az összekötöttség fokára érzéketlen egzakt algoritmust, mint például a felsorolásos (5.3.1 fejezet) vagy a

változóeliminációs (5.3.2 fejezet) eljárás. Ezek számítási igénye azonban (annak exponenciális volta miatt) a gyakorlatban csak játék-alkalmazásokban elfogadható. A hálót valamilyen eljárással átalakíthatjuk egy polifa struktúrába. Ebben az alfejezetben ilyen eljárásokra mutatunk egyszerű példákat, és az 5.4 alfejezetben tárgyalunk egy összetettebb algoritmust. Egy további lehetőség, hogy lemondunk az egzakt következtetésről és valamilyen sztochasztikus szimulációs (Monte-Carlo) eljárást alkalmazunk. Ezek bemutatásával és hátterükkel az 5.5 alfejezet foglalkozik.

Összevonós eljárások A háló csomópontjai közül bizonyosakat összevonunk úgynevezett megacsomópontokba amelyek az általuk tartalmazott eredeti csomópontok együttes feltételes eloszlását tartalmazzák. Az 5.5 ábra egy ilyen esetet illusztrál. Bár



5.5. ábra. Nem fa struktúrájú háló és a belőle összevonással létrehozott, megacsomópontokat tartalmazó transzformált háló.

a következtetés feladata ebben az új struktúrában már végrehajtható a lineáris időben, az eredeti hálóhoz viszonyítva ez nem áll fenn, hiszen az összevont csomópontok értékkészlete a tartalmazott csomópontok értékkészletének Descartes-szorzata.

Feltételezéses eljárások A *feltételezéses eljárások* az előző megközelítésnek éppen az ellenkezőjét alkalmazzák, itt az egyetlen, de (az egyes csomópontok szintjén) bonyolultabb másodlagos struktúra helyett több, de egyszerűbb hálót hozunk létre, amelyekben már egyszerűen végrehajtható a következtetés. Az itt bemutatott *vágóhalmaz feltételezésen alapuló eljárás* a következő fő lépésekből áll:

- Keressük meg a csomópontok egy olyan halmazát, amely konfigurációját rögzítve a háló struktúrája egy fává redukálódik (az ilyen csomópont-halmazt nevezzük *vágóhalmaznak*, jelöljük ezt C-vel).
- A vágóhalmaz lehetséges konfigurációival származtatott hálók mindegyikén végezzük el a következtetést: $P(X|E \cup \{C = c\})$.
- A keresett valószínűség a fenti valószínűségeknek a vágóhalmaz egyes konfigurációinak valószínűségével súlyozott átlagaként adódik:

$$P(X|E) = \sum_{\forall c} P(X|E \cup \{C = c\})P(C = c|E) \tag{5.20}$$

Ennél az eljárásnál is látható, hogy a következtetés exponenciális volta nem kerülhető meg: általános esetben a kiértékelendő fák száma a vágóhalmaz méretében exponenciális (egészen pontosan a vágóhalmaz elemei értékkészletei Descartes-szorzatának számossága).

Adott pontosságú vágóhalmaz feltételezés. Ennél a módszernél lehetőség van a számítási komplexitás csökkentésére –természetesen a pontosság rontása, és az egzaktságról való lemondás árán–: megtehetjük, hogy csak adott összvalószínűségű alhálóra végezzük el a következtetést és az ezekből adódó eredményt korrigáljuk a figyelembe vett hálók összvalószínűségével. Ha tehát a válasz pontosságának például 0,1-en belül kell lennie, annyi alhálót kell figyelembe vennünk, amennyinek az összvalószínűsége a 0,9-et meghaladja.

5.4. A PPTC-következtetés

A PPTC (*Probability Propagation in Trees of Clusters*) eljárás során az eredeti Bayes-hálóból egy másodlagos struktúrát hozunk létre, amely a változók eloszlását (illetve potenciálisan az evidenciákat is) már egy olyan formában tárolja, amely egyszerű eljárásokkal, közvetlenül felhasználható a lekérdezések megválaszolására. Az eljárás két fő részből áll: először a gráfstruktúrát transzformáljuk egy úgynevezett *klikkfába*; majd az eredeti háló csomópontjainak feltételes eloszlásait, illetve az evidenciákat visszük be a klikkfa csomópontjaiba.
Az eljárást először Lauritzen és Spiegelhalter publikálták [2]. Ebben a fejezetben Huang és Darwiche metodológiai cikkének [1] főbb elemeit ismertetjük.

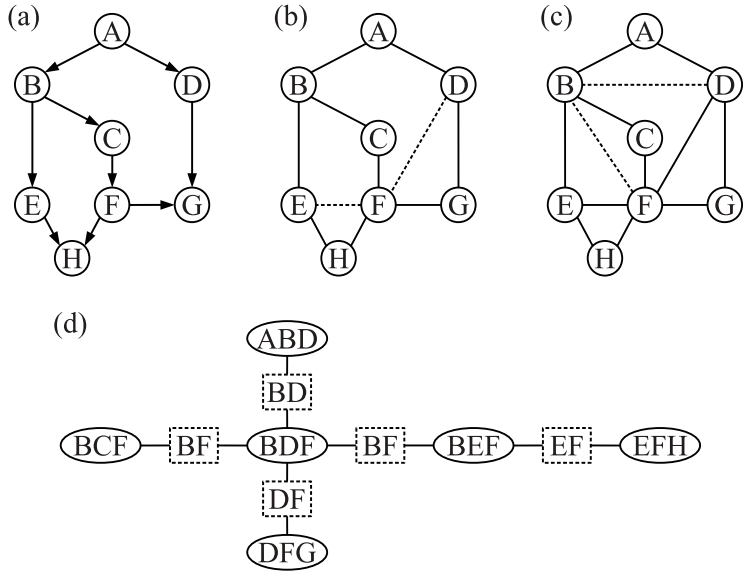
5.4.1. Klikkfa konstruálása

A PPTC-eljárás első részeként az eredeti struktúrából egy másodlagos struktúrát hozunk létre. A következőkben ennek a lépéseit vesszük sorra. Az 5.6 ábra ennek a folyamatnak a lépéseit illusztrálja.

Morális gráf formára alakítás Első lépésként a DAG-ot annak morális gráfjává alakítjuk, azaz összekötjük az egyes csomópontok szüleit egymással, majd töröljük az élek irányítottságát.

Háromszögesített gráf A kapott morális gráfot *háromszögesíteni* kell, azaz minden, 3-nál hosszabb körben kell lennie 2, a körben nem szomszédos, összekötött csomópontnak. A háromszögesítés megvalósítására az 5.2 algoritmus szolgálhat.
Mint látható, G_M hatékony háromszögesítéséhez minden csomóponthoz két számértéket kell számon tartani: az általa esetlegesen okozandó él-hozzáadások számát és a hozzá tartozó klikk súlyát. Ezeket az értékeket azonban nem kell minden csomóponttörléskor a teljes hálóra újraszámolni, elég csak az utoljára törölt V csomópont szomszédaira újraszámolni őket.

Klikkek azonosítása A háromszögesített gráfban azonosítani kell a maximális klikkeket (teljesen összekötött csomóponthalmazokat), ezek fogják a klikkfa csomópontjait alkotni. Bár vannak általános eljárások tetszőleges gráfban való klikk-keresésre, ebben a lépésben felhasználhatjuk, hogy (1) minden a gráfban lévő klikk a háromszögesítés során jött létre; és (2) egy ilyen klikk nem lehet egy korábban létrehozott másik klikk részhalmaza. Ebből következően elég, ha a háromszögesítés során feljegyezzük azokat



5.6. ábra. A másodlagos struktúra létrehozásának lépései a PPT- algoritmusban: az eredeti Bayes-háló struktúra (a), a morális gráf (b), a háromszögesített gráf (c), a klikkeket (folytonos vonal) és szeparáló halmazokat (szaggatott vonal) tartalmazó klikkfa (d)

Algorithm 5.2 Morális gráf G_M háromszögesítése

1: Készítsünk G_M-ről egy másolatot: G'_M
2: **while** $G'_M \neq \varnothing$ **do**
3: Válasszuk ki G'_M-ből a V csúcsot a 7. sortól leírt kritérium szerint.
4: V és szomszédai egy úgynevezett *klikket* alkotnak. Kössük össze az összes csomópontpárt a klikkben, és az újonnan hozzáadott éleket húzzuk be G_M-ben is.
5: Töröljük V-t G'_M-ből.
6: G_M az újonnan hozzáadott élekkel így már háromszögesítve van.

7: A 3. sorban alkalmazott kritérium:
8: Egy csomópont *súlya* annak értékkészletének számossága.
9: G'_M csomópontjai közül válasszuk azt, amelyik a 4 sorban a lehető legkevesebb él gráfhoz adását indukálja.
10: Ha több ilyen is van, válasszuk a legkisebb súlyút.

a klikkeket, amelyek az eddig lementettek egyikének sem részhalmazai: ez a lista meg fog egyezni a klikkek listájával.

Optimális klikkfa építése Az így előállított klikkekből ezután ki kell alakítani a klikkfát, mégpedig úgy, hogy lépésről-lépésre összekötögetjük a klikkeket, amíg a fa az összeset nem tartalmazza. A klikkek összekötése formálisan úgy történik, hogy közéjük a gráfban egy úgynevezett *szeparáló halmazt* (angol terminológiával sepset-et) illesztünk a gráfban. A sepset ugyanúgy az eredeti háló csomópontjainak egy halmazát

tartalmazza, mint a klikkek, nevezetesen azokat, amelyek a hozzá kapcsolódó mindkét klikkben jelen vannak.
Az, hogy a klikkeket milyen módon kötjük össze egy fában, a teljesítmény miatt lényeges. Az 5.3 algoritmus egy olyan kiválasztási kritériumot tartalmaz, amely biztosítja, hogy a klikkfa a következtetés számítási komplexitása szempontjából optimális legyen.

Algorithm 5.3 Optimális klikkfa építése

1: Kezdjünk n darab különálló, egy csomópontos gráffal (a $c \in \mathcal{C}$ klikkekből) és egy $\mathcal{S}$ üres halmazzal.
2: **for each** (X, Y) **in** $(X, Y) : X, Y \in \mathcal{C}, X \neq Y$ **do**
3: Állítsuk elő az S_{XY} sepset-et, amely X-re és Y-ra hivatkozik szomszédaiként és $X \cap Y$ csomópontokat tartalmazza.
4: Adjuk hozzá S_{XY}-t $\mathcal{S}$-hez.
5: **end for each**
6: **repeat**
7: Válasszunk ki egy S_{XY} sepset-et $\mathcal{S}$-ből, a 11. sortól kezdődően leírt kritérium szerint.
8: Töröljük S_{XY}-t $\mathcal{S}$-ből.
9: Illesszük be S_{XY}-t az általa hivatkozott X és Y klikkek közé, feltéve, hogy X és Y külön fákban vannak.
10: **until** Beillesztettünk $n - 1$ sepset-et.

11: A 7. sorban alkalmazott kritérium.
12: S_{XY} sepset *tömege* az általa tartalmazott csomópontok száma, azaz $|X \cap Y|$
13: S_{XY} sepset *költsége* X és Y klikkek súlyának összege, ahol egy klikk *súlya* az őt alkotó csomópontok súlyainak szorzata (lásd 5.2 algoritmus, 8. sor).
14: Az $S_{XY} \in \mathcal{S}$ sepset-ek közül válasszuk azt, amelyiknek a legnagyobb a tömege.
15: Ha több ilyen is van, akkor ezek közül válaszuk azt, amelyiknek a legkisebb a költsége.

Üres sepset-ek A fenti algoritmusban olyan sepset-ek is előfordulhattak, amelyek nem tartalmaztak egy csomópontot sem. Bár a legtöbb esetben ezek a sepset-ek nem jelennek meg a végső struktúrában, ha az eredeti háló nem volt teljesen összekötve, a klikkfa is tartalmazni fog ilyen sepset-et (egészen pontosan annyit, ahány a különálló komponensek összekötéséhez kell, vagyis ha az eredeti gráf K darab komponensből állt, akkor $K - 1$-et).

5.4.2. Valószínűségek terjesztése a klikkfában

Ha a fenti módon előállt a klikkfa struktúrája, a klikkek és sepset-ek által tárolt valószínűségi eloszlásokba be kell vinni az eredeti Bayes-háló által reprezentált értékeket. Ez a következő fő lépésekkel történik.

Inicializáció A csomópontok (vagyis klikkek és sepsetek) által tárolt valószínűségi potenciálok inicializálása az 5.4 algoritmus szerint történik.

Algorithm 5.4 Klikkfa potenciáljának inicializálása

1: **for each** X **in** $\mathcal{C} \cup \mathcal{S}$ **do**
2: $\quad \phi_X \leftarrow 1$
3: **end for each**
4: **for each** V **in** U **do**
5: $\quad$ Keressünk egy X klikket, hogy $V \cup Pa(V) \subseteq X$
6: $\quad \phi_X \leftarrow \phi_X P(V|Pa(V))$
7: **end for each**

Vagyis először minden klikk és sepset potenciálját „feltöltjük" 1-esekkel, majd pedig minden BN csomóponthoz keresünk egy olyan klikket, amely tartalmazza annak családját, és annak potenciáljába pontonkénti szorzással bevisszük a változó feltételes valószínűségi eloszlását, $P(V|Pa(V))$-t. Ezután a klikkfára igaz lesz a következő:

$$\frac{\prod_{i=1}^{N} \phi_{X_i}}{\prod_{j=1}^{N-1} \phi_{S_j}} = \frac{\prod_{k=1}^{Q} P(V_k|\mathbf{C}_{V_k})}{1} = P(U) \tag{5.21}$$

(ahol N a klikkek, Q a BN csomópontjainak száma, ϕ_H pedig a H halmaz feletti potenciál), vagyis a klikkfa globálisan konzisztens lesz.

Evidencia bevitele. Ha bizonyos változókról rendelkezésre állnak evidenciák is, akkor ezek is könnyen bevihetők a rendszerbe: ha λ_V jelöli a V változóval kapcsolatos ismereteinket leíró valószínűségi potenciált, akkor elég egy megfelelő (azaz V-t tartalmazó) X klikkbe bevinni ezt is egy pontonkénti szorzással:

$$\phi_X \leftarrow \phi_X \lambda_V \tag{5.22}$$

A valószínűségek globális terjesztése. A valószínűség-terjesztés célja, hogy a klikkfát lokálisan is konzisztenssé tegye, vagyis annak csomópontjai valóban az általuk tartalmazott változók együttes eloszlását tartalmazzák. Ezt a műveletet egy sor, szomszédos klikkek közötti úgynevezett üzenetküldéssel valósíthatjuk meg. Egy X-ből S_{XY}-on keresztül Y-ba küldött üzenet után R_{XY} konzisztens lesz X-szel (vagyis $\phi_{S_{XY}} = \sum_{X\setminus(X\cap Y)} \phi_X$), miközben a $P(U) = \frac{\prod_i \phi_{X_i}}{\prod_j \phi_{S_j}}$ egyenlet sem sérül.
A globális valószínűség-terjesztés során minden klikk minden egyes szomszédjához pontosan egy üzenetet fog küldeni, és a teljes folyamat végére a klikkfa lokálisan konzisztens lesz. Egyetlen üzenetküldés az alábbi két lépésben történik:

Projekció. Tároljuk S_{XY} régi tábláját, és rendeljünk hozzá egy újat:

$$\phi'_{S_{XY}} \leftarrow \phi_{S_{XY}} \tag{5.23}$$

$$\phi_{S_{XY}} \leftarrow \sum_{X\setminus S_{XY}} \phi_X \tag{5.24}$$

Abszorpció. Rendeljünk Y-hoz egy új táblát S_{XY} régi és új táblájának felhasználásával:

$$\phi_Y \leftarrow \phi_Y \frac{\phi_{S_XY}}{\phi'_{S_{XY}}}. \tag{5.25}$$

Belátható, hogy $\phi'_{S_{XY}}$ csak ott vehet fel 0 értéket, ahol $\phi_{S_{XY}}$ is. Az ilyen esetekben vegyük úgy, hogy $\frac{0}{0} = 0$

A teljes valószínűség terjesztés során egy tetszőlegesen választott X klikkből indulunk, majd ebből történik meg a $2(n-1)$ darab üzenetküldés. Ez két fázisban valósul meg, az első, úgynevezett bizonyíték-gyűjtési fázisban minden klikk X irányába küld üzenetet, a második, úgynevezett bizonyíték-elosztási fázisban (X-szel kezdve) minden klikk X-től távolabbi szomszédainak küld üzenetet, az 5.5 rekurzív algoritmus szerint. Ebben minden egyes klikkhez egy jelzőbitet ($b(X)$) rendelünk, amelynek segítségével számon tartjuk, hogy az adott fázisban mely klikkek küldtek már üzenetet és melyek nem.

Algorithm 5.5 Globális valószínűség-terjesztés

1: Válasszunk egy tetszőleges klikket: X.

2: **for each** C **in** $\mathcal{C}$ **do** $b(C) \leftarrow$ false
3: **end for each**
4: **call** COLLECT-EVIDENCE(X)

5: **for each** C **in** $\mathcal{C}$ **do** $b(C) \leftarrow$ false
6: **end for each**
7: **call** DISTRIBUTE-EVIDENCE(X)

8: **function** COLLECT-EVIDENCE(klikk X)
9: $b(X) \leftarrow$ true
10: **for each** Y **in** X szomszédai **do**
11: **if** $b(Y) =$ false **then call** COLLECT-EVIDENCE(Y)
12: **end for each**
13: Küldjünk üzenetet X-ből abba a csomópontba, amely ezt a függvényhívást indította.

14: **function** DISTRIBUTE-EVIDENCE(klikk X)
15: $b(X) \leftarrow$ true
16: **for each** Y **in** X szomszédai **do**
17: **if** $b(Y) =$ false **then**
18: Küldjünk üzenetet X-ből Y-ba.
19: **call** COLLECT-EVIDENCE(Y)
20: **end for each**

Normalizálás. Ha a fentiek során bizonyítékok bevitelére is sor került, az egyes csomópontok nem feltétel nélküli, hanem az evidenciákkal együtt vett eloszlásokat fognak tartalmazni, vagyis például $P(V, e)$-t. Ahhoz, hogy ezekből megkapjuk az általunk keresett $P(V|e)$-ket, szükséges még a klikkek és csomópontok tábláinak normalizálása is.

5.4.3. Következtetési esetek

Ha a fenti módon a klikkfát lokálisan és globálisan is konzisztenssé tettük, elkezdhetünk benne következtetni. Az alábbiakban röviden vázoljuk az alapvető következtetési esetek megoldását.

Egyetlen változó marginálisa. A lokális konzisztenciának megfelelően minden klikk és sepset táblája egy valós többváltozós marginálist tartalmaz. Ennek megfelelően, ha egyetlen változóra vagyunk kíváncsiak, elég keresnünk egy (lehetőleg minél kevesebb változót tartalmazó) csomópontot, és annak táblájából marginalizálással előállíthatjuk a keresett valószínűségi eloszlást.

Több változó marginálisa. Ha létezik olyan klikkfa csomópont, amely tartalmazza az összes lekérdezés-változót, a marginalizálás itt is végrehajtható. Ha viszont nincs olyan klikk, amely teljes egészében lefedné a célváltozók halmazát, a láncszabályt kell alkalmaznunk.
Egy másik lehetőség a globális konzisztencia, vagyis a

$$P(U) = \frac{\prod_i \phi_{X_i}}{\prod_j \phi_{S_j}} \tag{5.26}$$

egyenlet kihasználása. Ezt felhasználva a keresett valószínűségek szintén előállíthatók (bár a szerző tapasztalata szerint jóval kevésbé hatékonyan, mint a láncszabály alkalmazásával).

5.5. Közelítő következtetés sztochasztikus szimulációval

Sztochasztikus szimulációs eljárás alkalmazására akkor van szükség, amikor az egzakt következtetési eljárások - azok számítási költségei miatt - nem alkalmazhatók. A Monte-Carlo-eljárások alapfeladata a következő: adott egy valószínűségi változótól függő mennyiség, amelynek keressük a valószínűségi változó eloszlása szerinti várhatóértékét. Mivel a teljes eloszlás nem kezelhető, mintavételezzük azt, és a mintából számított statisztikával becsüljük a keresett mennyiséget. Képlettel kifejezve:

$$E_{P(X)}(f(X)) \approx \hat{E}_{P(X)}(f(X)) = \frac{1}{N}\sum_{i=1}^{N} f(X_i). \tag{5.27}$$

A mi konkrét esetünkben, a Bayes-hálókban való prediktív következtetés esetén, a mintavételezendő eloszlás a háló által reprezentált együttes eloszlás feltéve a bizonyítékokat, a keresett mennyiség pedig a célváltozó(k) marginális eloszlása. Az 5.3.1 alfejezetben bemutatott felsorolásos módszer a mintavételezendő eloszlás kimerítő felsorolásos feldolgozásának felel meg.
Mivel ez a naiv megközelítés a legritkább esetben vezet csak eredményre, ebben az alfejezetben először olyan egyszerűbb eljárásokat mutatunk be, amelyek magát a háló eloszlását mintavételezik; majd a *Markov-láncos Monte-Carlo* (Markov Chain Monte Carlo, MCMC) módszerek rövid elméleti áttekintése és a Bayes-hálós következtetésre való alkalmazásának bemutatása következik.

5.5.1. Mintagenerálás „üres" hálóból

A mintavételezésen alapuló minden következtetési eljárás a háló (és esetleg az evidenciák) alapján generál egy mintahalmazt, amely alapján már kiszámíthatunk egy statisztikát, amellyel a kérdéses valószínűséget becsülhetjük. A legegyszerűbb esetben az evidenciákat figyelmen kívül hagyva generálunk egy mintahalmazt a háló által leírt együttes eloszlásból.

Egyetlen minta generálása Bayes-hálóból Az eljárás alapeleme egyetlen minta generálása, amely a következő lépésekből áll:

1. Vegyük a csomópontok egy topologikus sorrendjét. Kezdetben minden csomópont álljon behelyettesített érték nélkül.

2. Ebben a sorrendben minden egyes csomópontra sorsoljunk ki egy értéket a háló által leírt feltételes valószínűségi eloszlásából, feltéve a szülők aktuális értékeit ($P(X|Pa(X))$). A csomópontnak adjuk értékül a kisorsolt értéket.

3. Ha végiglépdeltünk az összes csomóponton, a nekik adott értékek alkotják a kisorsolt konfigurációt.

5.5.2. Elutasító mintavételezés

A fenti eljárás kellő számú ismétlésével kapott minta már egyszerűen felhasználható egy adott (hálóbeli) esemény adott evidenciák melletti valószínűségének megbecslésére: ez egyszerűen a céleseménnyel és az evidenciákkal is kompatibilis minták számának aránya lesz az evidenciákkal kompatibilis esetek számához képest.
Az úgynevezett *elutasító mintavételezés* a fentieknek megfelelően még két lépést alkalmaz a generált mintán:

1. Kiszűri („elutasítja") a mintából azokat, amelyek nem kompatibilisek az evidenciákkal.

2. A fennmaradó mintán a kérdéses feltételes valószínűség a célkonfigurációval kompatibilis minták számának aránya a teljes (már megszűrt) mintaszámhoz viszonyítva.

A módszer fő előnye az egyszerűsége, hátránya viszont, hogy a becslés során valóban felhasználható minták száma az evidenciák mennyiségével exponenciálisan csökken.

5.5.3. Valószínűségi súlyozás

Ez az algoritmus az előző módosítása annak érdekében, hogy elkerülhessük a becslésben fel nem használható (az evidenciákkal inkompatibilis) minták generálását. Ennek eléréséhez az alap mintagenerálási lépést annyiban módosítjuk, hogy az evidencia-változók értékeit rögzítjük, és csak a fennmaradóknak sorsolunk. Az így nyert mintákat azonban súlyoznunk kell: míg korábban minden egyes minta 1 súllyal szerepelt, most ezt a súlyt meg kell szorozni az evidencia-változók feltételes valószínűségeivel. Ez az eljárás

annyiban különbözik az elutasító mintavételezéstől, hogy az evidenciákkal inkompatibilis mintákat már azok generálásakor „eldobjuk": egy-egy evidencia-változóra eső lépésben a kompatibilis sorsolások aránya az összeshez képest $P(E|Pa(E))$ lesz, ezért kell azzal súlyozni a mintát.
Az előző eljáráshoz hasonlóan a becsült valószínűség a lekérdezéssel kompatibilis minták és az összes minta összsúlyának arányaként áll elő.
Ez az eljárás hatékonyabb, mint az *elutasító mintavételezés*, mert figyelembe veszi az összes generált mintát. Az azonban még mindig problémát jelenthet hatékonysági szempontból, hogy előfordulhat, hogy a teljes minta összsúlyának egy nagyon nagy részét a minták egy kis hányada képviseli. Azok a minták ugyanis, amelyek valószínűségileg kevéssé illeszkednek csak a bizonyítékékhoz, kis súlyokat kapnak, így hatásuk a végső becslés szempontjából sem lesz jelentős.
A továbbiakban bemutatandó Gibbs-mintavételező ezt a gyengeséget orvosolja oly módon, hogy a mintavételezés során végig olyan konfigurációkat sorsol, amelyek az adott szituációban valószínűek. Mielőtt a Gibbs-mintavételező részletezésére rátérnénk, áttekintjük az annak alapjául szolgáló Monte-Carlo-eljárásokat.

5.6. A Monte-Carlo-eljárások áttekintése

A Monte-Carlo-eljárások alapötlete, hogy egy eloszlás feletti integrált vagy (az e fejezetben vizsgált diszkrét esetben) összegzést annak analitikus megoldása helyett valamilyen mintavételezési eljárással közelítik.
A kiszámítandó mennyiség:

$$\overline{f} = E_{\pi(X)}[f(X)], \tag{5.28}$$

amelynek közelítésére a következő lépéseket alkalmazzuk:

- Generáljunk egy $\{X_i\}_{i=1}^N$ *független, azonos eloszlású* (*independent, identically distributed - i.i.d.*) mintahalmazt $\pi(X)$ mintavételezésével.
- A minta alapján számoljuk ki az $\hat{f}_N = \sum_{i=1}^N f(X_i)$ közelítést.
- Adjunk valamiféle megbízhatósági becslést az $|\overline{f} - \hat{f}_N|$ valós és becsült érték közötti eltérésre.

Bayes-hálókban való következtetés esetén a mintavételezendő $\pi(X)$ eloszlás a háló által reprezentált $P(\mathcal{U})$ együttes eloszlás, esetleg ennek az evidenciák szerinti feltételes értéke: $P(\mathcal{U}|E)$. A keresett $\overline{f}$ érték a célváltozók marginális eloszlása, esetleg azok egy konfigurációjának valószínűsége.
A becslés megbízhatóságáról a következő eredmények léteznek:

- A *nagy számok erős törvénye* alapján az $\hat{f}_N$ becslés erősen konzisztens, azaz:

$$P(\lim_{N\to\infty} \hat{f}_N = \overline{f}) = 1. \tag{5.29}$$

- A *központi határeloszlás tétele* szerint $\hat{f}_N$ standardizáltja aszimptotikusan Gauss-eloszlású lesz:

$$\frac{\hat{f}_N - \overline{f}}{\sigma_N} \to N(0,1), \text{ ha } N \to \infty, \text{ ahol } \sigma_N = \frac{Var(f(X))}{\sqrt{N}} \tag{5.30}$$

- Ha, mint a mi esetünkben, $f(X)$ korlátos, akkor a közelítés konvergenciájának sebességére további becslések adhatók a Hoeffding- és a Bernstein-egyenlőtlenségek alapján (specifikusan, ha $0 \leq f(X) \leq 1$):

$$p(|\hat{f}_N - \overline{f}| \geq \epsilon) \leq 2\exp(-2\epsilon^2 N) \leq \delta \tag{5.31}$$

$$E[|\hat{f}_N - \overline{f}|] \leq \sqrt{\frac{c_0}{N}} \tag{5.32}$$

5.6.1. Fontossági mintavételezés

Abban az esetben alkalmazhatjuk a fontossági mintavételezést, ha a $\pi(X)$ eloszlás nem mintavételezhető hatékonyan, de rendelkezésre áll egy megegyező tartójú $q(X)$ eloszlás, amelyen ez megtehető. Ebben az esetben a $q(X)$-ből vett mintákat a következő képlettel használhatjuk fel $\overline{f}$ közelítésére:

$$\overline{f} = E_{\pi(X)}[f(X)] = E_{q(X)}[\frac{f(X)\pi(X)}{q(X)}] \tag{5.33}$$

vagyis az $\hat{f}_N$ közelítés az alábbi szerint számítható:

$$\hat{f}_N = \frac{\frac{1}{N}\sum_{t=1}^{N} w(X_t)f(X_t)}{\frac{1}{N}\sum_{t=1}^{N} w(X_t)} = \frac{1}{N}\sum_{t=1}^{N} w^*(X_t)f(X_t), \tag{5.34}$$

ahol $w(X_t)$ és $w^*(X_t)$ a fontossági súlyok:

$$w(X_t) = \frac{\pi(X_t)}{q(X_t)}, w^* = \frac{w(X_t)}{\frac{1}{N}\sum_{t=1}^{N} w(X_t)} \tag{5.35}$$

5.6.2. Markov-láncok

5.3. *definíció* (Markov-lánc). Az $\mathcal{X} = \{X_0, X_1, \dots\}$ valószínűségi változók sorozata egy *(elsőrendű) Markov-lánc*, ha $P(X_t|X_{t-1}, \dots X_0) = P(X_t|X_{t-1})$ A Markov-lánc *homogén*, ha $P(X_t|X_{t-1})$ nem függ t-től.

A fenti X_t változókat általában a Markov-lánc *állapotainak* tekintjük, amelyeket az egymás után felvesz. Ahhoz, hogy egy folyamat Markov-lánc lehessen - mint az a definícióból látszik - az kell, hogy állapotai a múltbeli állapotoktól csak a megelőzőn keresztül függjenek. A t paraméternek gyakran valamiféle időbeli interpretációt tulajdonítunk.
A továbbiakban az X_t állapotokat természetes számokkal $(i, j, \dots)$ jelöljük, és mivel feltesszük, hogy a vizsgált Markov-láncok homogének (vagyis idő-invariánsok), az i állapotból j-be való átlépés valószínűségét $p_{ij} = P(X_t = j|X_{t-1} = i)$-vel jelöljük. A p_{ij} értékekből alkotott *egylépéses állapot-átmeneti mátrix* $P = P[p_{ij}]$. Az n-lépéses állapot-átmeneti mátrix P n-edik hatványa: $P^{(n)} = P^n$.

5.4. *definíció* (Invariáns eloszlás). A p'^{inv} eloszlást a χ homogén Markov-lánc invariáns eloszlásának nevezzük, ha $p'^{inv} = p'^{inv}P$, ahol P χ állapotátmenet-mátrixa. (Következésképpen ha $p^{(0)} = p'^{inv}$, akkor $\forall t : p^{(t)} = p'^{inv}$.)

Markov-láncok tulajdonságai

Az alábbiakban sorra vesszük a Markov-láncok legfontosabb tulajdonságait, és a hozzájuk kapcsolódó alapvető tételeket.

5.5. *definíció* (Stabilitás). Az $\mathcal{X}$ Markov-lánc *stabil*, ha a $\lim_{t\to\infty} p(X_t) = p^{(\infty)}$ létezik, valóban egy eloszlás és független a $p(X_0)$ kiindulási eloszlástól. $p^{(\infty)}$-t *határérték-* vagy *egyensúlyi eloszlásnak* nevezzük.

5.6. *definíció* (Irreducibilitás). A diszkrét, véges állapotú $\mathcal{X}$ Markov-lánc *irreducibilis*, ha bármely (i, j) állapotpárra létezik olyan $n_{ij} > 0$ lépésszám, hogy a $p_{ij}^{(n_{ij})}$ állapotátmenet-valószínűség nagyobb mint 0.

5.7. *definíció* (Aperiodicitás). A diszkrét, véges állapotú $\mathcal{X}$ Markov-lánc *aperiodikus*, ha van olyan i állapot (irreducibilitást feltételezve ez bármely i-re igaz), hogy létezik $n_i > 0$ lépésszám, hogy bármely $n \geq n_i$-re $p_{ii}^{(n)} > 0$

5.3. tétel. Ha egy diszkrét, véges állapotú $\mathcal{X}$ Markov-lánc irreducibilis és aperiodikus, akkor $\mathcal{X}$ stabil, és létezik egyetlen invariáns eloszlása, amely $\mathcal{X}$ határérték-eloszlása (vagyis p'^{∞} az egyetlen, nemnegatív megoldása $p'^{\infty} = p'^{\infty}P$-nek és $p^{(\infty)}$ eloszlás).

A stabil Markov-láncok esetén ezt a határérték-eloszlást (p^{∞}, p^{inv}) $\pi(X)$-szel jelöljük.

5.4. tétel. Ha a diszkrét, véges állapotú $\mathcal{X}$ Markov-lánc stabil és $\overline{f} = E_{\pi(X)}[f(X)] < \infty$, akkor $P(\lim_{N\to\infty} \hat{f}_N = \overline{f} = 1)$, ahol $\hat{f}_N = \frac{1}{N}\sum_{t=1}^{N} f(X_t)$

5.8. *definíció* (Ergodicitás). A diszkrét, véges állapotú $\mathcal{X}$ Markov-lánc *geometrikusan ergodikus*, ha létezik olyan $0 \leq \lambda < 1$ és $V(.) > 1$ függvény, hogy

$$\forall i : \sum_j |p_{ij}^{(t)} - \pi_j| \leq V(i)\lambda^t \tag{5.36}$$

A legkisebb ilyen λ-t a *konvergencia sebességének* hívjuk.

5.5. tétel. Legyen $\mathcal{X}$ egy diszkrét, véges állapotú, geometrikusan ergodikus (vagyis egyúttal stabil) Markov-lánc, annak invariáns eloszlásából $\pi(X)$ indítva, f pedig egy valós értékű függvény, amelyre igaz, hogy $E_\pi[f(X)^{2+\varepsilon}] \leq \infty$ valamely $\varepsilon > 0$-ra. Ekkor $\overline{f} = E_\pi[f(X)]$, és $\sigma^2 = Var_\pi(f(X))$ jelölések mellett $\hat{f}_N = \frac{1}{N}\sum_{t=1}^{N} f(X_t)$-re:

$$\tau^2 = \sigma^2 + 2\sum_{k=1}^{\infty} E_\pi[(f(X_0) - \overline{f})(f(X_k) - \overline{f})] \tag{5.37}$$

létezik, nemnegatív és véges, valamint

$$\sqrt{N}\frac{\hat{f}_N - \overline{f}}{\tau} \xrightarrow{d} N(0,1),\ ha\ N \to \infty \tag{5.38}$$

5.6.3. A Metropolis-Hastings-algoritmus

Jelölje $\pi(X)$ a normalizálatlan, szigorúan pozitív ($\forall i : \pi_i = \pi(X = i) > 0$) céleloszlást az $S = 0, 1, \ldots K$ halmaz felett. Legyen Q, egy átmenet-valószínűségi mátrix ($Q\underline{1} = \underline{1}$), az úgynevezett javasolt eloszlás, úgy hogy $q_{ij} > 0$ akkor és csak akkor, ha $q_{ji} > 0$. Definiáljuk a χ Markov-láncot a P állapotátmenet-mátrixszal a következőképpen:

$$\forall i \neq j : p_{ij} = q_{ij}\min(1, \frac{\pi_j q_{ji}}{\pi_i q_{ij}}), \tag{5.39}$$

$\frac{0}{0} = 0$-t feltéve és $p_{ii} = 1 - \sum_{j \neq i} p_{ij}$ definícióval. Fontos látni, hogy csak a π céleloszlás elemeinek arányaira van szükségünk, ami jól illeszkedik a bayesi analízis során előforduló normalizálatlan eloszlások alkalmazásához.
A definiált Markov-lánc stacionárius eloszlása $\pi(X)$ lesz, ami könnyen belátható, látván, hogy a részletes egyensúly-feltétel teljesül. Az $i = j$ és a $q_{ij} = q_{ji} = 0$ esetek triviálisan teljesülnek. Ha $i \neq j$ és $q_{ij} > 0$, akkor feltéve, hogy $\pi_i q_{ij} \geq \pi_j q_{ji}$:

$$\pi_i p_{ij} = \pi_i \frac{\pi_j q_{ji}}{\pi_i q_{ij}} = \pi_j q_{ji} = \pi_j p_{ji} \tag{5.40}$$

Ha Q irreducibilis, akkor P is az lesz, és ez hasonlóképpen igaz az aperiodikusságra is. Következésképpen ha találunk egy olyan Q javaslati eloszlást, amely irreducibilis és aperiodikus, akkor egy adott $\pi(X)$ céleloszláshoz a fenti konstrukció egy stabil és reverzibilis Markov-láncot definiál, amelynek (invariáns) határeloszlása $\pi(X)$ lesz.
A Metropolis-Hastings-algoritmus alkalmazása tehát az alábbi fő lépésekből áll:

0. (Konstruáljuk meg a poszterior eloszlás egy P^S közelítését; ezt az MCMC inicializálására és később esetlegesen ellenőrzésre használhatjuk.)

1. Konstruáljunk egy irreducibilis és aperiodikus Q javaslati eloszlást, specifikusan a doménhez.

2. Sorsoljunk egy x_0 kiinduló állapotot P^S-ből.

3. $t = 1, 2, \ldots$-ra:
   Sorsoljunk egy x^* jelölt állapotot Q-ból, feltéve x_t-t.
   Számítsuk ki az α elfogadási valószínűséget az x_t-ből x^*-ba történő lépésre:

$$\alpha(x_t, x^*) = \min(1, \frac{\pi_{x^*} q_{x^*, x_t}}{\pi_{x_t} q_{x_t, x^*}}) \tag{5.41}$$

   Legyen $x_{t+1} = x^*$ $\alpha(x_t, x^*)$ valószínűséggel, különben legyen $x_{t+1} = x_t$.

4. Folytassuk ezt, amíg az a lánc nem konvergál, és el nem éri az előírt konfidenciát.

5. (Értékeljük ki a konvergencia sebességét és javítsunk a hatékonyságon Q újratervezésével; ezután lépjünk vissza a 2. lépésre.)

6. (Vessük össze a módszert P^S fontossági újramintavételezésén alapuló alap módszerrel; ezután lépjünk vissza az 1. lépésre.)

A Metropolis-Hastings algoritmus alesetei

Az ismertetett Metropolis-Hastings algoritmus a legáltalánosabb MCMC-eljárás; bizonyos paramétereinek különféle beállításaival számos, a szakirodalomban számon tartott speciális esetét kapjuk. Ezek közül a következőket említjük meg:

Metropolis-algoritmus. Akkor áll elő, ha a Q állapotátmeneti mátrix szimmetrikus.

Véletlen bolyongás Metropolis-algoritmus. Ha Q csak a jelenlegi és a javasolt állapot között értelmezett valamilyen távolságtól függ: $q(x^*|x_t) = q(|x^* - x_t|)$.

Független mintavételező. Ha Q nem függ a jelenlegi állapottól: $q(x^*|x_t) = q(x^*)$.

Gibbs-mintavételező. Ha Q olyan, hogy az állapotvektornak egyszerre mindig csak egy elemét változtatja meg, annak a többitől való feltételes eloszlása szerint, 1 elfogadási valószínűséggel.

5.6.4. Következtetés Bayes-hálókban Gibbs-mintavételezéssel

A Gibbs-mitavételező az alábbi lépéseket követi.

1. Kisorsolunk egy konfigurációt, amely kompatibilis a bizonyítékokkal.
2. Rögzítjük az evidencia-változók értékét.
3. A fennmaradó változókon periodikusan újrasorsoljuk az aktuális változó értékét a többi változó (praktikusan az aktuális változó Markov-takarója) mint feltétel szerint.
4. Az előző lépésben kapott minden egyes konfiguráció egy-egy eleme lesz a teljes mintahalmaznak.

5.7. Függelék: A következtetés komplexitása Bayes-hálókban

Egy Bayes-háló „mérete” Ahhoz, hogy beszélni tudjunk egy eljárás tár- és időigényéről, meg kell tudnunk határozni a bemenet (esetünkben az adott Bayes-háló) hosszát. Egy Bayes-háló leírásához annak struktúráját és a csomópontokhoz tartozó feltételes eloszlásokat kell megadnunk. Ezt megtehetjük úgy, hogy minden csomópontra megadjuk annak szüleit, és a CPT-bejegyzéseit. Mivel ez utóbbiak száma a család (a gyermek és szülei) tagjai értékkészlete Descartes-szorzatának számossága, vagyis a szülők számában exponenciális, a strukturális információ mértékét elhanyagolhatjuk, és tekinthetjük a Bayes-hálót meghatározó paraméterek számát az összes CPT-bejegyzés számának.
Mint az előző fejezetekben láttuk, a következtetés komplexitása egy naiv, az összes lehetséges változó-konfigurációt felsoroló eljárás esetében arányos a konfigurációk számával, azaz a komplexitás a változók számában exponenciális, vagyis valós alkalmazások esetén a következtetés kivitelezhetetlen. Az alábbiakban bemutatjuk, hogy az

ismert *3SAT* probléma visszavezethető a Bayes-hálókban való következtetésre, amiből következik, hogy az NP-nehéz. A következőkben néhány, az algoritmikus problémák nehézségével és nehézségük összehasonlításával kapcsolatos alapfogalmat tekintünk át.

5.9. *definíció* (P). $\mathrm{P} = \cup_{k\geq 1}\mathrm{TIME}(n^k)$, ahol TIME(.) a Turing-gépek által a bemenetnek a paraméterben megadott függvényével arányos időben felismerhető (megoldható) nyelvek (problémák) osztálya.

P tehát azon nyelvek (problémák) halmaza, amelyek a bemenet hosszának valamilyen polinomiális függvényével arányos időben felismerhetők (eldönthetők). Általánosan kijelenthető, hogy a P-be tartozó problémák „könnyen", „hatékonyan" megoldhatók, kezelhetők.

5.10. *definíció* (NP). $\mathrm{NP} = \cup_{k\geq 1}\mathrm{NTIME}(n^k)$, ahol NTIME(.) a nemdeterminisztikus Turing-gépek által a bemenetnek a paraméterben megadott függvényével arányos időben felismerhető (megoldható) nyelvek (problémák) osztálya.

Az NP-beli problémák tehát olyanok, amelyek esetén egy megoldásról polinom időben eldönthető, hogy az valóban helyes-e, arról azonban, hogy ezt a megoldást hogyan és mennyi idő alatt lehet megtalálni, már nem állítunk semmit. Mint a következőkben látni fogjuk, számos „nehéz" NP-beli probléma létezik, ha pedig ezeket vissza tudjuk vezetni egy másik (az általunk éppen vizsgált) problémára, akkor ily módon beláthatjuk (vagy legalábbis elfogadhatjuk), hogy a mi adott problémánk is ugyanolyan nehéz. A visszavezetés fogalmának precízebb definíciója az alábbi.

5.11. *definíció* (Karp-redukció). Az f leképezés az L_1 nyelv *Karp-redukciója* az L_2 nyelvre, ha $\forall x$ szóra $x \in L_1$ pontosan akkor teljesül, ha $f(x) \in L_2$, és f polinom időben számítható. Jelölése: $L_1 \prec L_2$.

5.12. *definíció* (NP-teljesség). Egy L nyelv *NP-teljes*, ha eleme NP-nek és bármely $L' \in NP$ nyelvhez létezik $L' \prec L$ Karp-redukció.

5.13. *definíció.* Egy L nyelv *NP-nehéz*, ha létezik hozzá egy L' NP-teljes nyelv, hogy $L' \prec L$.

Belátható, hogy ha egy adott L_1 NP-teljes nyelvhez létezik $L_2 \in$ NP nyelv, amelyre $L_1 \prec L_2$, akkor L_2 is NP-teljes.
Összefoglalva: ha tudunk adni egy polinomidejű algoritmust, amely egy ismerten NP-nehéz problémát a mi problémaosztályunk egy elemére vezet vissza, azzal bizonyíthatjuk, hogy a mi problémánk is NP-nehéz.

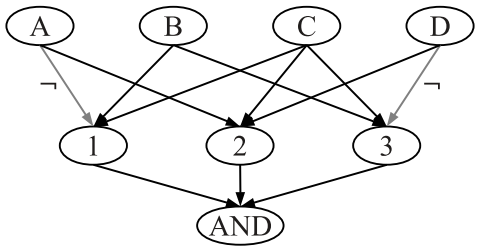
5.7.1. A 3SAT probléma visszavezetése a Bayes-hálóban való következtetésre

Konjunktív normál formának (conjunctive normal form, CNF) hívjuk az olyan logikai kifejezéseket, amelyek literálok diszjunkcióinak („vagy" kapcsolatainak) konjunkciójaként („és" kapcsolataként) állnak elő, azaz például $(x_1 \vee \overline{x}_2 \vee x_3) \wedge (x_2 \vee x_4) \wedge \ldots$ alakúak. Egy *n-CNF* olyan CNF melynek egy-egy tagjában maximum n literál szerepel.
A 3-CNF-ek kielégíthetőségét vizsgáló *3SAT* problémáról (amely tehát azt vizsgálja, hogy egy adott 3-CNF változói behelyettesíthetőek-e úgy, hogy a kifejezés igaz legyen) ismert, hogy NP-teljes. Ha tehát tudunk olyan eljárást adni, amely minden 3SAT

problémát visszavezet egy Bayes-hálóban való következtetésre, akkor azzal igazoljuk, hogy a Bayes-hálókban való következtetés NP-nehéz.
Vegyük a következő kifejezést:

$$(\overline{A} \vee B \vee C) \wedge (A \vee C \vee D) \wedge (B \vee C \vee \overline{D}). \tag{5.42}$$

Ebből az 5.7 ábrán látható struktúrát konstruáljuk, vagyis:



5.7. ábra. A 3SAT *probléma visszavezetése egy Bayes-háló struktúrára*

- A háló minden csomópontja az $\{igaz, hamis\}$ tartományból kaphat értéket.
- A kifejezésben szereplő változók a háló felső sorának csomópontjai lesznek, $\{0{,}5, 0{,}5\}$ a priori eloszlással.
- Az egyes tagoknak egy-egy, a középső szinten elhelyezkedő csomópont felel meg, amely feltételes eloszlása az adott tag által leírt kifejezésnek lesz megfelelő (például az „1”-es csomópont 1 valószínűséggel *igaz* lesz, ha B és C közül legalább az egyik igaz, vagy ha A hamis, minden más esetben 1 valószínűséggel *hamis* lesz).
- A háló „AND” csomópontja az előző szinthez hasonlóan fogja össze a középső csomópontok értékét „és” kapcsolattal.

Belátható, hogy ebben a hálóban az $E = \{\text{AND} = igaz\}$ bizonyíték mellett azon behelyettesítéseknek lesz nem nulla a valószínűsége, amelyek a kifejezést igazzá teszik. Ebből már látható, hogy a leképezett 3SAT kifejezést kielégítő kifejezés keresése visszavezethető a konstruált hálóban való következtetésre. Ezzel beláttuk, hogy a következtetés általános esetben NP-nehéz.

Irodalomjegyzék

[1] Cecil Huang and Adnan Darwiche. *Inference in belief networks: A procedural guide. International Journal of Approximate Reasoning, 15(3):225-263, 1996.*

[2] S. L. Lauritzen and D. J. *Spiegelhalter. Local computations with probabilities on graphical structures and their application to expert systems. Journal of the Royal Statistical Society. Series B (Methodological), 50(2):157-224, 1988.*

[3] Stuart J. Russel and Peter Norvig. *Mesterséges intelligencia modern megközelítésben. Panem, 2005.*

6. fejezet

Döntéstámogatás: optimális döntés, szekvenciális döntések, az információ értéke

6.1. Szekvenciális döntési folyamatok

6.1.1. Optimális döntés

Egy döntési helyzetben tipikus feladat a rendelkezésre álló információ alapján egy, vagy több kritériumnak megfelelően a legjobb lehetőség kiválasztása a felmerülő opciók közül. Ilyen egyszerű helyzet lehet egy üdülés célpontjának kiválasztása, miközben adott a költségkeret vagy a kiindulási hely, és vannak bizonyos preferenciáink (például tengerpart, pálmafák).
Egy döntéselméleti helyzet leírásához definiálni kell egy rendszer **állapotait** ($s \in S, states$), az egyes állapotokban elérhető **lépéseket** ($a \in A, actions$), és meg kell határozni az egyes állapotok **hasznosságát** ($U(s), utility$). Egyes állapotokból a megfelelő lépés kiválasztásával lehetséges az átlépés más állapotokba. A hasznosságfüggvény az állapotok felett definiált valós függvény, amely teljesíti a racionális döntéshozóra vonatkozó **hasznossági axiómákat**. A hasznosságfüggvény valós függvény, ezért az axiómák többségét a definícióból fakadóan kielégíti. Ezek az axiómák a sorrendezhetőség, a tranzitivitás, a folytonosság, a monotonitás. A döntéselméleti helyzeteket sztohasztikus jellegük miatt szokás szerencsejátéknak is nevez. A szerencsejátékokra vonatkozó további axiómák a következőek:

- Helyettesíthetőség - Ha egy olyan szerencsejátékot játszunk, amelyben p valószínűséggel az egyik, $1-p$ valószínűséggel a másik állapotba jutunk, és van két állapot, melynek hasznossága azonos (s_i, s_j), akkor azok a szerencsejátékban felcserélhetőek:

$$U(s_i) = U(s_j) \Rightarrow \exists p [p, s_i; 1-p, s_k] \sim [p, s_j; 1-p, s_k].$$

- Felbonthatóság – Ha egy olyan összetett szerencsejátékot játszunk, amelyben az első szerencsejáték egyik kimenete egy másik szerencsejáték, akkor az ekvivalens

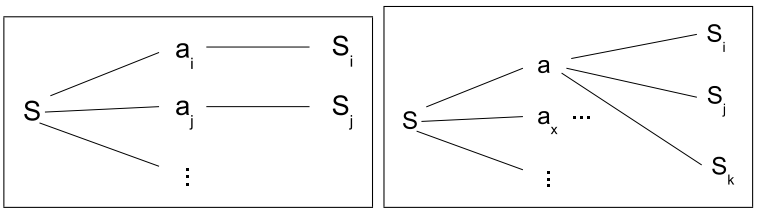
egy három kimenetű összetett szerencsejátékkal, függetlenül az állapotok hasznosságától, illetve p és q értékétől:

$$[p, s_i; 1-p, [q, s_j; 1-q, s_k]] \sim [p, s_i; (1-p)q, s_j; (1-p)(1-q), s_k]$$

Legegyszerűbb esetben a lépések eredménye **determinisztikus** (6.1 ábra), így az optimális döntés minden s állapotban az az a lépés, amely az elérhető legnagyobb hasznosságú állapothoz vezet (6.1 egyenlet).

$$U(a^*|s) = max_{a \in A} U(a|s). \tag{6.1}$$

Abban az esetben viszont, amikor a rendszer valamilyen bizonytalanságot tartalmaz, a lépések kimenetele **nemdeterminisztikus** (6.1 ábra), így egy eloszlást definiálhatunk a lépések kimenete felett: s_i állapotban, a lépés mellett annak valószínűsége, hogy s_j állapotba jutunk $P(s_j|a, s_i)$.



6.1. ábra. Bal oldal: Determinisztikus döntési helyzetben a_i választása mellett s_i állapotba jutunk. Jobb oldal: Nemdeterminisztikus döntési helyzetben a választása mellett $P(s_i|a, s)$ valószínűséggel s_i állapotba jutunk.

A racionális döntéshozóra vonatkozó axiómákból következik, hogy nemdeterminisztikus esetben az optimális döntés mindig azt az a^* lépést jelenti, amely maximalizálja a várható hasznosságot (maximal expected utility, MEU):

$$MEU(a^*|s) \doteq \max_{a \in A} EU(a|s) = \max_{a \in A} \sum_{s_i \in S} U(s_i) P(s_i|a, s), s \in S. \tag{6.2}$$

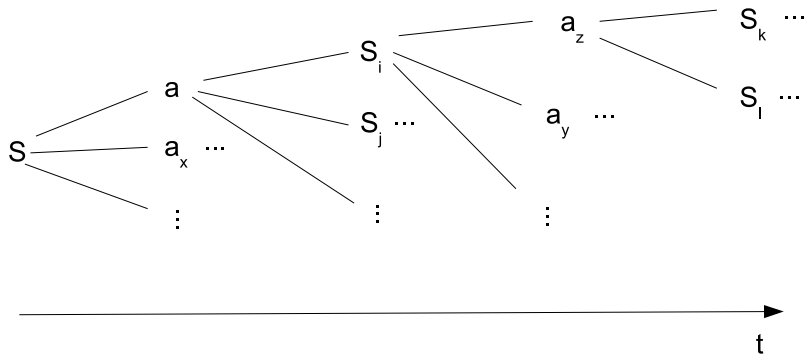
6.1.2. Szekvenciális döntés

Az előző fejezetben az **egylépéses**(myopic) esetet vizsgáltuk, bizonyos esetekben azonban előfordulhatnak **szekvenciális** döntési helyzetek (6.2 ábra). Például egy körutazás esetén minden érintett helyhez rendelhetünk hasznosságot, a pillanatnyi tartózkodási helytől pedig függ a másnap elérhető helyek halmaza. A bizonytalanságot a rendszerben a megbízhatatlan idegenvezető jelentheti.
Egylépéses esetben egy állapot hasznosságát definíció szerint az $U(s)$ hasznosságfüggvény határozta meg. Szekvenciális esetben azonban egy állapot hasznosságát befolyásolja a belőle elérhető további állapotok hasznossága is. Az egyes állapotokra vonatkozó $U(s)$ hasznosságfüggvények felhasználásával a t diszkrét időpontban s állapot hasznossága rekurzív képlettel írható le:

$$U^t(s) = U(s) + \sum_{s_i \in S} U^{t+1}(s_i) P(s_i|a, s). \tag{6.3}$$

Tehát $U(s)$ az a hasznosság, amelyet s meglátogatása ér, $U^t(s)$ jelenti azt a hasznosságot, amely figyelembe veszi az s állapotból elérhető teljes döntési gráfot és az abból számolt várhatóértéket. Az egyenlet második tagja egy várhatóérték az s állapotból elérhető állapotok hasznossága felett. A várható maximális hasznosság ennek megfelelően módosul:

$$MEU(a^*|s) = \max_{a \in A} \sum_{s_i \in S} U^t(s_i) P(s_i|a, s), s \in S. \tag{6.4}$$



6.2. ábra. Szekvenciális döntés

Feltételezve, hogy minden állapot elérhető minden állapotból, és egy állapotban többször is tartózkodhatunk, a 6.3 egyenlet számítása n **véges** lépést feltételezve dinamikus programozással $O(n|A||S|)$ ideig tart. Az utolsó lépésben egyszerűen adódik, mekkora az egyes állapotok hasznossága ($U(s)$), majd az $n-1$ lépésben a hasznosságok már a 6.3 képlettel kiszámolhatóak, egészen a kezdő lépésig. Természetesen ez a módszer nem alkalmazható végtelen számú lépés esetén.

Végtelen lépésszám esetén ($n \to \infty$) a 6.4 egyenlet nem alkalmazható, mert a nyelő csomópontoktól eltekintve a maximális várható hasznosság minden állapotra végtelennek adódik. Egy lehetséges megoldás a jövőbeni jutalmak leszámított értékével történő számítás, ekkor a jövőbeni jutalom egy $0 < \gamma \leq 1$ együtthatóval megszorzott értékével számolunk:

$$MEU'(a^*|s) = \max_{a \in A} \sum_{s_i \in S} \gamma U^t(s_i) P(s_i|a, s), s \in S. \tag{6.5}$$

A 6.5 egyenletet lépésenként kibontva $[a_0^*, a_1^*, \ldots, a_i^*, \ldots]$ optimális akciók egy sorozatát hajtjuk végre. Ha feltételezzük, hogy egy állapotból maximum k másik állapotba lehet eljutni, $\gamma < 1$, az állapotváltások valószínűségének maximuma P_{max} és az $U(s)$ függvény maximuma U_{max}, akkor a következő kifejezésre jutunk:

$$\begin{aligned} MEU'(a_0^*, a_1^*, \ldots, a_i^*, \ldots | s) &\leq U(s) + \gamma k U_{max} P_{max} + \\ &+ \gamma^2 k U_{max} P_{max} + \ldots + \gamma^i k U_{max} P_{max} + \ldots \\ &= U(s) + k U_{max} P_{max} \sum_{i=0}^{\infty} \gamma^i \\ &= U(s) + \frac{k U_{max} P_{max}}{1-\gamma}. \end{aligned}$$

A 6.6 egyenletből látszik, hogy a $\gamma < 1$ feltétel, és a leszámított számítás esetén a maximális várható hasznosság felülről becsülhető.

Markov döntési folyamatok

Az előzőekben bevezetett döntéselméleti formalizmus meghatározó jellemzője a **Markov-tulajdonság**: annak valószínűsége, hogy a folyamat t időpillanatban s_i állapotba kerül, csak a folyamat $t-1$ időpillanatban felvett s állapottól függ, ha az ismert ($P(s_i|a,s)$). Tehát s állapot ismeretében a korábbi állapotok ismerete nem szükséges az állapotátmenetek valószínűségének meghatározásához.
A Markov döntési folyamat az eddigiektől kicsit eltérő, azonban azokkal teljesen ekvivalens, szekvenciális döntési folyamatokat leíró igen elterjedt formalizmus. A Markov döntési folyamat definiálja az S_0 kezdőállapotot, a $T(s,a,s')$ állapotátmenet modellt (transition) és az $R(s)$ jutalomfüggvényt (reward). A $T(s,a,s')$ állapotátmenet függvény ekvivalens a korábban definiált $P(s_i|a,s)$ feltételes valószínűséggel, míg az $R(s)$ jutalomfüggvénynek az $U(s)$ hasznosságfüggvény felel meg. Markov döntési folyamatokkal kapcsolatban szokás beszélni az úgynevezett eljárásmódról (policy), mely a döntéshozónak minden állapotra meghatározza, hogy adott állapotban melyik lépést válassza. Optimális eljárásmódnak (optimal policy) nevezzük azt az eljárásmódot, amely a maximális várhatóértékű lépést adja. A korábban bevezetett fogalmakkal az optimális eljárásmód azt jelenti, hogy a döntéshozó minden lépésben az a^* lépést választja (véges lépésszám esetén a 6.4 , végtelen lépésszám esetén a 6.6 egyenlet alapján).

6.1.3. Az információ értéke

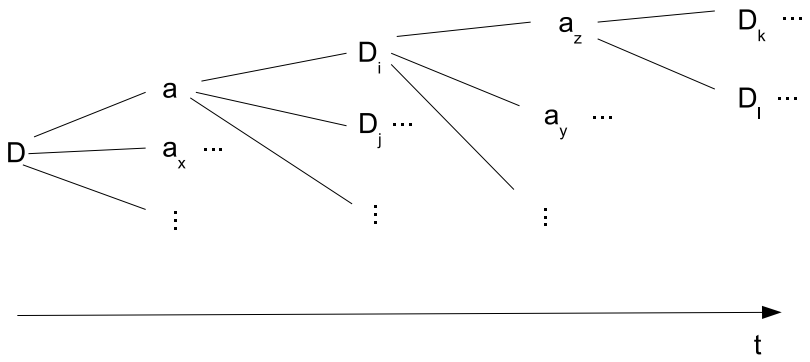
Az előzőekben a szekvenciális döntéseket úgy modelleztük, hogy minden döntési lépés után a rendszer állapotot vált. Azonban sok esetben döntési opció a szekvenciális döntési sorozatból történő kilépés, a megállás. Ebben az esetben természetesen merül fel az igény a jövőben rendelkezésünkre álló adat értékének ismeretére. Ezt írja le a tökéletes információ értéke (value of perfect information, VPI), nagyon hasonlóan a 6.2 egyenlethez (lásd 6.3 ábra):

$$MEU(a^*|d) = \max_{a \in A} EU(a|d) = \max_{a \in A} \sum_{d_i \in \mathcal{D}} U(d_i) P(d_i|a,d), d \in \mathcal{D}, \tag{6.6}$$

ahol d a rendelkezésre álló adatot jelenti, $\mathcal{D}$ minden lehetséges adathalmaz összessége, míg d_i az az adat, amelyhez akkor jutunk, ha az a lépést választjuk és ezzel például folytatjuk az adatgyűjtést. A 6.6 egyenlet valójában csak annyiban tér el a 6.2 egyenlettől, hogy az s-t a d helyettesíti, vagyis jelenleg a rendszer állapotát (a világról gyűjtött

információt) a rendelkezésre álló adattal írjuk le. Fontos megjegyezni, hogy $d \subset d_i$. Ha rendelkezésre állna d_i információ, akkor a várható hasznosság a következőképpen alakulna:

$$MEU(a^*|d_i) = \max_{a \in A} EU(a|d_i) = \max_{a \in A} \sum_{d_j \in \mathcal{D}} U(d_j) P(d_j|a, d_i), d_i \in \mathcal{D}. \qquad (6.7)$$



6.3. ábra. Az információ értéke

A 6.6 és a 6.7 kifejezések által definiált értékek eltérése adná meg a kívánt mennyiséget, vagyis annak a plusz információnak a hasznosságát, amelyet $d_i \setminus d$ ismerete jelent. Azonban d_i nem áll rendelkezésünkre, ezért jelölje D a jövőbeli adatot reprezentáló valószínűségi változót, így VPI a MEU jövőbeli várhatóértékének és a MEU-nak a különbsége:

$$VPI_d(D_j) = \left[\sum_i P(D_j = d_{ij}|d) MEU(a^*_{D_j,d}|D_j = d_{ij}, d) \right] - MEU(a^*_d|d), \qquad (6.8)$$

ahol a^*_d a d adat ismeretében a legjobb döntés (lépés).

A VPI tulajdonságai

1. A VPI nem vehet fel negatív értéket,

   $$VPI_d(D_i) \geq 0,$$

   szemléletesen azért, mert az újonnan megszerzett információtól mindig el lehet tekinteni. A MEU érték maximum képzés eredménye, ezért ha bármely újabb d_i információ mellett a MEU kisebb értéket venne fel, a maximum képzés miatt d_i-t üres adatnak kell feltételezni, hogy a legjobb eredményt kapjuk. Így a VPI_d érték nullának adódik.

2. Könnyen belátható, hogy az információ értéke az adatok beérkezésének sorrendjétől független. Az információ értéke számolható a következőképpen:

$$VPI_d(D_i, D_j) = VPI_d(D_i) + VPI_{d,D_i}(D_j) = VPI_d(D_j) + VPI_{d,D_j}(D_i). \quad (6.9)$$

3. Nem igaz azonban, hogy az információ értékének képzése additív

$$VPI_d(D_i, D_j) \neq VPI_d(D_i) + VPI_d(D_j),$$

hiszen például abban az esetben, ha a D_i és D_j valószínűségi változók azonos eloszlásúak $VPI_d(D_i, D_j) = VPI_d(D_i) = VPI_d(D_j)$.

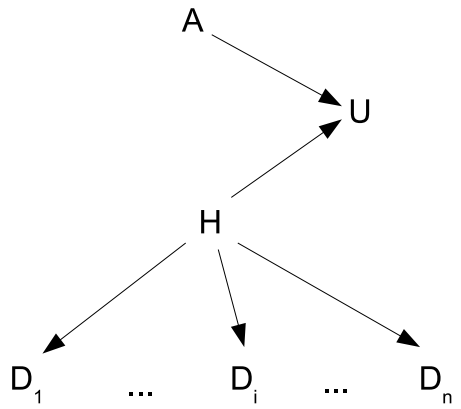
Az információ értékének közelítése több megfigyelés esetén

Az információ értékével kapcsolatban eddig egy olyan esetet vizsgáltunk, amikor minden lépést megelőzően egyetlen valószínűségi változó jövőbeli várhatóértékét számítottuk ki. A gyakorlati esetek többségében a VPI számításához a 6.1.3 fejezetben tárgyalt egyváltozós módszert használják. Előfordulhat azonban olyan eset, amikor a döntési lépést megelőzően több valószínűségi változó értékét is meg kell becsülni. A 6.8 és a 6.9 egyenletekből könnyen levezethető, hogy a VPI kiszámításához szükséges idő több D_i valószínűségi változó esetében azok n számával exponenciálisan arányos.

Ebben a fejezetben több megfigyelés együttes információértékét becsüljük az eddig ismertetettektől eltérő döntéselméleti modell mellett (6.4 ábra). Tegyük fel, hogy a $D_i, i = 1, \ldots, n$ valószínűségi változók függetlenek, és a döntési lépést egy A bináris valószínűségi változóval modellezünk. Tegyük fel továbbá, hogy egy ismert eloszlású, bináris H valószínűségi változóval magasabb szinten tudjuk leírni a rendelkezésünkre álló $D_i, i = 1, \ldots, n$ adatot, vagyis ismertek a $P(D_i|H)$ feltételes eloszlások. Az A lépés és a H hipotézis közös hasznosságfüggvénye $U(A, H)$. Ezzel a modellel közelítő becslés adható a rendelkezésre álló változóhalmaz információértékére vonatkozóan lineáris időben.

Látni fogjuk, hogy a bizonyítás során nagyban kihasználjuk azokat az egyszerűsítéseket, amelyeket az A változó bináris volta és a szintén bináris H hipotézis változó jelent. Az utóbbi felfogható úgy, mint a rendelkezésre álló D_i adatok egy absztrakt, egyszerűsített leírása. Ez az egyszerűsített modell a többszörös megfigyelés információértékének lineáris időben történő kiszámításához szükséges, mégsem valóságtól elrugaszkodott példa: képzeljük el, hogy a D_i változók egy beteg különböző leleteit reprezentálják, míg a H változó azt a feltételezést, hogy a beteg a leletek alapján súlyos betegségben szenved. Ha az A döntés a műtét elrendelését jelenti, akkor az $U(A, H)$ hasznosság azt jellemzi, hogy a betegség esetleges megléte mellett mennyire kockázatos vagy hasznos a műtét, illetve annak elkerülése.

Ha felírjuk a H hipotézisre vonatkozó feltételes valószínűségek hányadosát (odds), akkor az a Bayes-szabálynak és a D_i valószínűségi változók függetlenségének köszönhetően átalakítható a következőképpen:



6.4. ábra. Információ értékének becslése több megfigyelés esetén

$$\begin{aligned} O(H|D_1,\ldots,D_n) &= \frac{P(H|D_1,\ldots,D_n)}{P(\neg H|D_1,\ldots,D_n)} \\ &= \frac{P(D_1|H)}{P(D_1|\neg H)}\cdots\frac{P(D_n|H)}{P(D_n|\neg H)}\frac{P(H)}{P(\neg H)} \\ &= O(H)\prod_{i=1}^{n}\lambda_i, \end{aligned} \tag{6.10}$$

ahol $\lambda_i = \frac{P(E_i|H)}{P(E_i|\neg H)}$ és $O(H) = \frac{P(H)}{P(\neg H)}$.
Legyen p^* a H hipotézis bekövetkezésének valószínűsége, amikor is indifferens a döntéshozó számára, hogy mely lépést választja, formálisan:

$$p^*U(H,A) + (1-p^*)U(\neg H,A) = p^*U(H,\neg A) + (1-p^*)U(\neg H,\neg A). \tag{6.11}$$

A 6.11 egyenletet átrendezve a p^* valószínűségre a következő érték adódik a hasznosságok ismeretében:

$$p^* = \frac{U(\neg H,\neg A) - U(\neg H,A)}{U(\neg H,\neg A) - U(\neg H,A) + U(H,A) - U(H,\neg A)}. \tag{6.12}$$

Mivel p^* valószínűség a döntési küszöb, a döntéshozó akkor választja A-t $\neg A$-val szemben, ha

$$P(H|D_1,\ldots,D_n) > p^*. \tag{6.13}$$

Ezt átírva a következőt kapjuk:

$$O(H|D_1,\ldots,D_n) > \frac{p^*}{1-p^*}. \tag{6.14}$$

A 6.10 egyenlet alapján a 6.14 kifejezés átírható

$$\prod_{i=1}^{n}\lambda_i > \frac{p^*}{1-p^*}/O(H). \tag{6.15}$$

Ha a 6.15 mindkét oldalának természetes alapú logaritmusát vesszük, akkor

$$\sum_{i=1}^{n} w_i > ln\frac{p^*}{1-p^*} - lnO(H), \tag{6.16}$$

ahol $w_i = ln\lambda_i$, így a W valószínűségi változó, mint w_i változók összege definiálható:

$$W \equiv \sum_{i=1}^{n} w_i, \tag{6.17}$$

és felírható a W változóhoz tartozó döntési küszöbérték:

$$W^* \equiv ln\frac{p^*}{1-p^*} - lnO(H), \tag{6.18}$$

vagyis a döntéshozó akkor dönt A lépés mellett, ha

$$W > W^*. \tag{6.19}$$

W valószínűségi változó a w_i független valószínűségi változók összege, ezért a centrális határeloszlás tétele alapján eloszlása normális és várhatóértéke a w_i változók várhatóértékének összege, míg szórása a w_i változók szórásának összege, így:

$$p(W|H) \sim N(E(W|H), Var(W|H)). \tag{6.20}$$

A 6.20 egyenlet alapján kiszámítható, hogy mi annak valószínűsége, hogy a W valószínűségi változó a küszöbérték felett lesz:

$$p(W > W^*|H) = \frac{1}{\sigma\sqrt{2\pi}} \int_{W^*}^{\infty} e^{\frac{-(t-\mu)^2}{2\sigma}} dt. \tag{6.21}$$

6.2. Megállási feladatok

Ahogy azt az előző fejezetben említettük, szekvenciális döntési helyzetben lehetséges lépés a leállás. Minden olyan esetben, amikor a továbblépés költséggel jár, bármely lépésben optimális döntés lehet a megállás.
Szekvenciális kiválasztási probléma esetében a döntéshozónak egy n hosszú, szekvenciálisan érkező $X_1, \ldots, X_n$ változó sorozatból ki kell választania a legnagyobbat úgy, hogy a be nem érkezett változókról semmilyen információja nincs, korlátozottan választhat a már beérkezettek közül, és a játék bizonyos variációiban n végtelen, vagy nincs róla információ. Az egyik legegyszerűbb megállási probléma az ún. Titkárnő probléma.

6.2.1. Titkárnő probléma

A megállási problémák alapfeladata a titkárnő probléma, amikor a munkáltatónak a legmegfelelőbb munkaerőt kell kiválasztania a pozícióra. A feladat a következő szabályokkal definiálható:

1. Csak egyetlen szabad állás van.

2. A jelentkezők száma, n, előre ismert.

3. Az interjúkat egymás után, egyesével bonyolítják le.

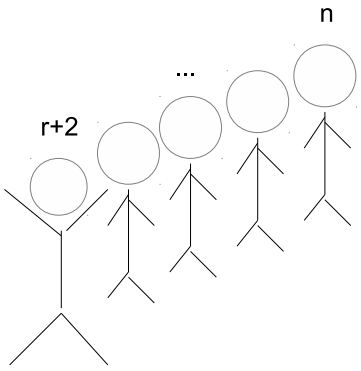
4. A jelentkezők meghallgatása véletlenszerű sorrendben történik, minden sorrend egyformán valószínű.

5. Az addig meghallgatott jelentkezők minden interjú után alkalmasságuk szerint egyértelműen rendezhetők.

6. Minden interjú után el kell dönteni, hogy a jelentkezőt felveszik-e, vagy sem. Ha egy jelentkezőt nem vesznek fel, nem lehet őt többé visszahívni.

7. A munkáltatónak csak a legalkalmasabb jelölt felel meg, minden más jelölt azonos mértékben alkalmatlan.

A döntéshozó igen nehéz helyzetben van, mert bár van információja a már elküldött jelentkezőkről, nem tudja őket visszahívni, azokról a jelentkezőkről pedig, akik még nem voltak interjún semmilyen információja nincs. A probléma megoldása a következő felismerésből adódik: a döntéshozó minden esetben csak a már meglévő információ alapján dönthet, vagyis érdemes megfelelő mennyiségű információt begyűjteni, hogy aztán ezek alapján a lehető legnagyobb valószínűséggel el lehessen dönteni egy jelentkezőről, hogy az a legjobb-e. A megoldásként adódó algoritmus:



6.5. ábra. Titkárnő probléma

1. Az első $r-1$ jelentkező meghallgatása után,

2. azt a jelentkezőt kell választani, amelyik jobb, mint az első $r-1$ jelentkező bármelyike.

Annak valószínűsége, hogy adott r mellett a fenti algoritmussal a legjobb jelentkezőt választjuk:

$$P_{opt}(r) = P(r \text{ mellett a legjobbat választjuk}) = \sum_{i=r}^{n} \frac{1}{n} \frac{r-1}{i-1}, \tag{6.22}$$

mivel az $\frac{r-1}{i-1}$ hányados annak a feltételes valószínűségét adja, hogy ha i a legjobb jelölt, akkor az előző $i-1$ jelentkező közül a legjobb az első $r-1$ jelentkező között van. Minden esetben a 6.22 kifejezést maximalizáló r-t kell választani.
Bizonyítható, hogy n növekedtével az optimális r tart n/e -hez, és annak a valószínűsége, hogy az algoritmus a legjobb jelentkezőt választja tart $1/e$-hez, ahol e az Euler-féle szám. Vagyis annak a valószínűsége, hogy megtaláljuk a legjobb jelöltet megközelítőleg $0,368$.
Az állítás bizonyításához először átalakítjuk a 6.22 kifejezést, majd belátjuk, hogy $n \to \infty$ esetén $P_{opt}(r) \to -xln(x)$, amelynek szélsőértéke könnyen meghatározható. Az első lépés a $P_{opt}(r)$ átalakítása:

$$\begin{aligned} P_{opt}(r) &= \sum_{i=r}^{n} \frac{1}{n}\frac{r-1}{i-1} \qquad (6.23) \\ &= \frac{r-1}{n}\sum_{i=r}^{n}\frac{1}{i-1} \\ &= \frac{r-1}{n}\sum_{i=r}^{n}\frac{n}{i-1}\frac{1}{n}. \qquad (6.24) \end{aligned}$$

Egy tetszőleges $f(\cdot)$ függvény bal oldali Riemann összege az $[a,b]$ intervallumon:

$$\Delta x(f(a)+f(a+\Delta x)+f(a+2\Delta x)+\ldots+f(b-\Delta x)) = \sum_{i=0}^{\frac{b-a}{\Delta x}} f(a+i\Delta x)\Delta x. \qquad (6.25)$$

A fenti egyenlet $f(t)=1/t$ függvény esetén $\Delta x = 1/n$ lépésközzel, ahol $a=\frac{r-1}{n}$ és $b=1$:

$$\begin{aligned} \sum_{i=0}^{\frac{b-a}{\Delta x}} f(a+i\Delta x)\Delta x &= \sum_{i=0}^{n-r+1}\frac{n}{(r-1+i)}\frac{1}{n} \\ &= \sum_{i=r-1}^{n}\frac{n}{i-1}\frac{1}{n}. \qquad (6.26) \end{aligned}$$

Tehát a 6.26 $n \to \infty$ esetén

$$\lim_{n\to\infty}\sum_{i=r-1}^{n}\frac{n}{i-1}\frac{1}{n} = \int_{\frac{r-1}{n}}^{1}\frac{1}{t}dt. \qquad (6.27)$$

A 6.27 határértéket visszaírva a 6.23 egyenletbe az $x=\frac{r-1}{n}$ helyettesítéssel:

$$\begin{aligned} \lim_{n\to\infty}\frac{r-1}{n}\sum_{i=1}^{n}\frac{n}{i-1}\frac{1}{n} &= x\int_{x}^{1}\frac{1}{t}dt \\ &= -xln(x). \qquad (6.28) \end{aligned}$$

Mivel $-xln(x)\frac{dx}{dt} = 1+ln(x)$, ami az $x=1/e$ helyen veszi fel a 0 értéket, bizonyítottuk az eredeti állítást.

6.2.2. A Googol játék

A Googol játék a megállási problémák egyik első verziója, amelyet Martin Gardner publikált 1960-ban. A Googol játékban ketten vesznek részt. Az egyik szereplő előre meghatározott n számú lapra felír általa választott, különböző egész számokat. A másik szereplő az eddigi döntéshozó helyzetében van: a lefordított lapok közül addig húz, amíg úgy nem gondolja, hogy a legnagyobb számot tartalmazó lapot tartja a kezében. Ez a feladatkiírás annyiban tér el a korábbitól a döntéshozó szemszögéből, hogy nem feltételezheti az egymás után érkező elemek függetlenségét.
Ha elfogadjuk a feltételezést, hogy az előre kiválasztott n szám együttes eloszlása leírható egy egyváltozós sűrűségfüggvénnyel, melynek egyetlen argumentuma az n szám maximuma

$$p(x_1, \ldots, x_n) = g(max\{x_1, \ldots, x_2\}), \tag{6.29}$$

vagyis ha az egyenlet által meghatározott értelemben az $\{X_1, \ldots, X_n\}$ számsorozat felcserélhető, akkor bizonyítható, hogy $n > 2$ esetén a Googol játék esetén r a következőképpen adódik:

$$\frac{1}{r} + \frac{1}{r+1} + \ldots + \frac{1}{n-1} < 1 < \frac{1}{r-1} + \frac{1}{r} + \ldots + \frac{1}{n-1}. \tag{6.30}$$

Az is belátható, hogy a Googol játékot játszó személy számára a következő két eset ekvivalens:

1. Nem tud semmit a lapokon található számokról (akár determinisztikusak is lehetnek)

2. A számok egyenletes eloszlásúak a $(0, \beta)$ intervallumon, ahol β ismeretlen.

6.2.3. Odds algoritmus

Az odds algoritmus több megállási probléma megoldását adja meg azáltal, hogy egy általánosabban megfogalmazott feladatot old meg: adott $I_1, \ldots, I_n$ indikátorváltozó sorozat, ahol I_j változó A_j esemény bekövetkezését mutatja. Az események egymás után, egyesével következnek be. A cél egy olyan módszert megadni, amely biztosítja, hogy a döntéshozó a legnagyobb valószínűséggel álljon meg az utolsó bekövetkező eseménynél. A $max_t\{P(I_t = 1, I_{t-1} = 0, \ldots, I_1 = 0)\}$ kifejezést maximalizáló index a τ leállási idő. Belátható, hogy ha I_j bekövetkezésének valószínűsége p_j és $o_j = p_j/(1 - p_j)$ (odds, arány), akkor τ az első olyan index, amelyre $I_\tau = 1$, $\tau > r_n$ és

$$r_n = \max\{1, \max\{1 \leq k \leq n : \sum_{j=k}^{n} o_i \geq 1\}\}. \tag{6.31}$$

Vagyis az r index pontosan akkor optimális, ahonnan kezdődően a hátralévő odds-ok összege először nagyobb, mint 1, vagy ha nincs ilyen index, akkor az $r = 1$ érték. Annak valószínűsége, hogy az eljárással az első „sikeres” eseménynél állunk meg:

$$P(I_\tau = 1, I_{t-1} = 0, \ldots, I_1 = 0) = \left(\prod_{j=r}^{n}(1 - p_j)\right)\left(\sum_{j=r}^{n} o_j\right). \tag{6.32}$$

Az odds algoritmussal megoldható a titkárnő probléma, ha az eredeti feladatot átfogalmazzuk a következőképpen: $I_k = 1$, ha a k sorszámú jelentkező jobb, mint a korábbiak ($X_k > X_i, \forall i < k$), így $P(I_k) = 1/k$, $o_k = 1/(k-1)$. Tehát az az optimális r, emylnél az $R = 1/(n-1) + 1/(n-2) + \ldots + 1/(r-1)$ összeg nagyobb lesz, mint 1. Ha $n \to \infty$, akkor $R \to 1/e$, ahogy korábban is láttuk.

6.2.4. Az odds algoritmus egy folytonos kiterjesztése

Időben többször bekövetkező független események között eltelt idő modellezésére használt valószínűségi változó eloszlása folytonos esetben exponenciális eloszlású, mivel a folytonos eloszlások közül ez az egyetlen örökifjú tulajdonságú

$$P(X_T > x_s + x_t | X_T > x_s) = P(X_T > x_t), \forall x_s, x_t > 0. \tag{6.33}$$

A 6.33 egyenlet szemléletesen annyit jelent, hogy ha egy teremben x_s ideje várakozunk, annak valószínűsége, hogy a további x_t időintervallumban nem fog villanykörte kiégni, nem függ x_s-től, vagyis nem függ attól, mennyi ideje várunk. Ha a villanykörték kiégését független eseményként kezeljük, akkor az utolsó villanykörte kiégése semmilyen hatással nincs a következő körte kiégésének bekövetkeztére, ezért a teremben a körték kiégése között eltelt idő exponenciális eloszlású valószínűségi változóval írható le. További példa lehet ilyen folyamatokra egy kevésbé forgalmas úton közlekedő autók közti távolság, vagy a beérkező telefonhívások közt eltelt idő.
Ha az események között eltelt idő λ paraméterű exponenciális eloszlású,

$$f(x;\lambda) = \left\{ \begin{array}{ll} \lambda e^{\lambda x} & \text{ha } x > 0, \\ 0 & \text{egyébként,} \end{array} \right\} \tag{6.34}$$

akkor annak valószínűsége, hogy τ időintervallumban az adott esemény k alkalommal fordul elő, λ paraméterű homogén Poisson-eloszlást követ:

$$P[(N(t+\tau) - N(t)) = k] = \frac{e^{-\lambda\tau}(\lambda\tau)^k}{k!} \qquad k = 0, 1, \ldots, \tag{6.35}$$

ahol $N(t)$ a t időpillanatig bekövetkezett események száma. Ha a λ intenzitás paraméter időben változhat, azaz $\lambda(t)$, akkor inhomogén Poisson-beszélünk.
Ha a $\lambda(t)$ paraméterű Poisson eloszlású valószínűségi változók „sikerességét” a $h(t)$ sűrűségfüggvény írja le, akkor a megállási probléma a következőképpen módosul: állítsuk meg a játékot a $[0, T]$ időintervallumban az utolsó sikeres eseménynél. A folytonos feladat a következőképpen oldható meg: a $[0, T]$ időintervallumot m részre osztva, annak valószínűsége, hogy a k sorszámú intervallumban legalább egy sikeres esemény következik be $p_k = \lambda(t_k)h(t_k)(t_k - t_{k-1}) + o(t_k - t_{k-1})$. Ha az intervallumok számát növelve, azok mérete tart a nullához, $(t_k - t_{k-1}) \to \infty$, akkor az intervallumban a bekövetkezés valószínűsége tart az intenzitás és a sikeresség valószínűségének szorzatához, $p_k \to \lambda(t_k)h(t_k)$. Ezek alapján a diszkrét esethez nagyon hasonló a megoldás:

$$\tau = \sup\left\{0, \sup\left\{0 \le t \le T : \int_t^T \lambda(u)h(u)du \ge 1\right\}\right\} \tag{6.36}$$

6.3. Többkarú rabló feladatok

A többkarú rabló probléma (multi-armed bandit problem, MAB) erőforrás allokációs probléma. Alapfeladata megfeleltethető a szerencsejátékos problémájának: a játékos k ún. félkarú rablóval (egy bizonyos fajta szerencsejáték-automatával) játszva szeretné maximalizálni a várható nyereményét. A játékos minden lépésben választ egy játékautomatát, melynek meghúzza a karját. Az a gép, amelyiknek meghúzza a karját, a gépre és annak pillanatnyi állapotára jellemző valószínűségi eloszlás szerint fizet jutalmat. A valós helyzettől a többkarú rabló alapprobléma annyiban tér el, hogy a gépek működtetésének nincs költsége. A cél minden esetben az erőforrások optimális kihasználása: véges horizonton a begyűjtött jutalmak összegének maximalizálása, végtelen horizonton adott diszkontrátával vagy végtelen horizonton átlagban.
A többkarú rabló probléma k független karból/folyamatból/gépből és egy kontroller folyamatból áll. (E három fogalmat: kar, folyamat, gép a fejezetben mostantól felváltva használjuk.) Minden karhoz két véletlen folyamat tartozik $(X(0), X(1), \ldots,, R(X(0)), R(X(1)), \ldots,)$, ahol az $X(n)$ a kar állapota azután, hogy a kart n-szer működtettük, $R(X(n))$ pedig az $X(N)$ állapotért kapott jutalom. Egy gép állapota a következőképpen változik:

$$X(n) = f_{n-1}(X(0), \ldots, X(n-1), W(n-1)),$$

ahol $f(\cdot)$ adott és $W(n)$ egy ismert eloszlású, valós-értékű, független, azonos eloszlású valószínűségi változó sorozat, mely független $X(0)$-tól. Mivel az $f(\cdot)$ függvény determinisztikus, a $W(n)$ valószínűségi változó jelenti a véletlen faktort a gép működésében. A k-karú rabló probléma k karja független egymástól, a karokat egy kontroller/processzor folyamat működteti, minden diszkrét időpillanatban egy és csak egy kart választva ki. A kiválasztott folyamat állapotot vált, a többi folyamat állapota változatlan marad. A cél a várható jutalom maximalizálása. A többkarú rabló feladatok alkalmazásait és lehetséges medolgásait a következő alfejezetekben tárgyaljuk.

6.3.1. Alkalmazási területek

1. Szenzor menedzsment - egy egyszerűsített példa azt szemlélteti, hogy egy szenzorral hogyan keresünk egy célpontot, mely k lehetséges dobozok egyikében van. A szenzorunk képes érzékelni a célpontot, de csak bizonyos bizonytalansággal. Célunk egy előre meghatározott küszöbnél nagyobb bizonyosság elérése. Minden lépésben választhatunk, hogy melyik doboz által kibocsátott jelet mérjük meg a szenzor segítségével. A jutalom szerepét a szenzor által mutatott jelszint tölti be. A jelszint arányos annak valószínűségével, hogy a célpont az aktuális dobozban található. Azzal tehát, hogy a várható jutalmat növeljük, a célpontot keressük.

2. Online hírdetések kiválasztása - a hírdetést megjelenítő tartalomszolgáltató (például weboldal, mobilos alkalmazás) minden oldalmegjelenéskor kiválaszthatja a hírdetők által felkínált reklámok közül azokat, amelyek kikerülnek a felületre. A tartalomszolgáltató akkor jut bevételhez, ha az olvasó a hírdetésre kattint, ezért minden oldalmegjelenéskor azt a reklámot választja, amelyre a legnagyobb valószínűséggel kattintanak. A tartalomszolgáltató rendelkezésére állnak az eddigi megjelenések és kattintások, így minden oldalmegjelenéskor egy valószínűségi

változó írja le annak esélyét, hogy megjelnésekor egy reklámra kattintanak. A reklámok közül történő sorozatos választás egy többkarú rabló feladattal modellezhető.

3. Sorbanállási és ütemezési feladatok - a MAB-feladat felfogható egy egyetlen processzorból, és k feladatból álló rendszernek, amelyben minden egyes lépésben el kell dönteni, hogy a processzor mely feladatot hajtsa végre. Minden feladat végrehajtásáért jár egy jutalom, amely például a feladat sürgősségét tükrözi.

4. Klinikai kísérlettervezés - a gyógyszerkísérletekben a gyógyszerek hatóanyagainak megválasztása szintén megfeleltethető a MAB problémának. Itt az egyes karokat a hatóanyagok jelentik, míg a jutalmat a betegek állapota, esetleg túlélési rátája. A legnagyobb várható jutalomtól a legjobb hatóanyag kiválasztását reméljük.

6.3.2. Az optimális megoldás, előrefele következtetés

A MAB-probléma esetében a visszafele következtetés (backward induction) minden esetben optimális megoldást ad, viszont rendkívül számításigényes, ezért a gyakorlatban igen ritkán alkalmazzák.
Az előrefele következtetés legegyszerűbb formája az egylépéses (myopic) előretekintés, amely egyetlen lépésre előre maximalizálja a jutalmat. Ez a megoldás általában nem vezet optimális megoldáshoz. Az előrefele következtetés bonyolultabb típusa T lépésre előre számítja a várható jutalmat és ezt az értéket próbálja maximalizálni. Ezzel a megoldással az esetek nagy részében csupán szuboptimális megoldáshoz jutunk.
A T lépéses előretekintés kiterjesztésének esetében feltételezzük, hogy egy végrehajtási stratégia adott. Ezen ismert végrehajtási stratégia függvényében határozunk meg egy τ „leállási" időt, melyre a végrehajtási folyamat a maximális várható jutalmat adja. A végrehajtást csak a meghatározott leállási időpillanatig folytatjuk, az optimalizációt csupán erre kell végrehajtani.
Az előrefele következtetés a τ leállási idő számításával a következőképpen alakul:

1. Ki kell választani egy stratégiát. A stratégia ebben az esetben egyetlen gép működtetését jelenti.

2. Ki kell számítani egy τ leállási időt

3. A τ leállási időig követjük az 1. pontban választott stratégiát. τ után újból az 1. lépéssel folytatjuk.

Általános esetben a fent leírt stratégia sem vezet optimális megoldáshoz, azonban az alábbi feltételek mellett az algoritmus optimális a MAB problémára:

1. A kontroller folyamat egy időben csak egyetlen gépet üzemeltet; az üzemeltetett gép állapota nem befolyásolható, csak ki- és bekapcsolni lehet a gépet.

2. A nem működtetett gép nem vált állapotot.

3. A gépek függetlenek egymástól

4. A nem működtetett gépek nem adnak jutalmat.

A fent ismertetett algoritmus optimalitását, a megadott feltételek mellett szemléletesen úgy láthatjuk be, hogy minden alkalommal, amikor kiválasztunk egy gépet, majd azt τ-ig működtetjük, nem hozunk „visszafordíthatatlan" döntést. A többi gép állapota nem változik, vagyis nincs olyan jutalom, melyet hosszútávon az algoritmus ne tudna megszerezni, így az előrefele következtetés optimális megoldáshoz vezet.

6.3.3. Gittins index

Nem biztos, hogy egy folyamat t időpillanatig t-szer vált állapotot (hiszen üzemeltethetjük a többi folyamatot is), ezért i folyamat állapotát t-ben $N_i(t)$-vel jelöljük. Egy folyamatot az állapot és a jutalom sorozata írja le: $(X_i(N_i(t)), R_i(N_i(t)))$; $N_i(t) = 1, 2, \ldots, t; t = 1, 2, \ldots$ és $i = 1, 2, \ldots, k$. $U(t)$ vektor jelöli, hogy t időpillanatban melyik folyamatot üzemelteti a kontroll folyamat. $U(t) = (U_1(t), \ldots, U_k(t)), U(t)$ minden időpillanatban egyetlen komponensében sem nulla. A komponens indexe az adott pillanatban üzemeltetett folyamat indexét jelöli.
A MAB alapfeladata, hogy maximalizálja a következő kifejezést:

$$J = E\left[\sum_{t=0}^{\infty} \beta^t \sum_{i=1}^{k} R_i \left[X_i(N_i(t), U_i(t)) | X_1(N_1(0)), \ldots, X_k(N_k(0))\right]\right], \tag{6.37}$$

ahol $0 < \beta < 1$, vagyis a jutalom várható jelenértékét maximalizáljuk.
A Gittins-index a következő kifejezést takarja:

$$v_{x_i}(x_i(0)) = max_{0<\tau} \frac{E\left[\sum_{t=0}^{\tau-1} \beta^\tau R_i(X_i(t)|x_i(0))\right]}{E\left[\sum_{t=0}^{\tau-1} \beta^\tau |x_i(0))\right]}, \tag{6.38}$$

vagyis a Gittins-index azt jelenti, hogy minden egyes karra meghatározunk egy olyan τ megállási időt, amelyre nézve a fenti hányados maximális. Így a már tárgyalt előrefele következtetés algoritmusa a következőképpen alakul:

1. Minden folyamatra kiszámítjuk a Gittins-indexet, ezzel együtt minden folyamatra meghatározunk egy τ leállási időt.

2. Kiválasztjuk a maximális indexszel rendelkező folyamatot, majd az indexhez tartozó leállási ideig működtetjük. Leálláskor az 1. ponttal folytatjuk.

A fenti algoritmus markovi feltételezés mellett optimális megoldáshoz vezet.

7. fejezet

Orvosi döntéstámogatás

7.1. Egészségügyi adatok és nyilvántartó rendszerek

Az orvosi adatok kezelése, tárolása, az orvosi döntéshozatal folyamán történő optimális felhasználása mára önálló területté vált. Az orvosi döntési protokollok komplexitása, az elvégezhető vizsgálatok sokszínűsége, valamint a páciensek nyomon követésének igénye szükségessé tette a megfelelő informatikai háttér kialakítását. Az erre szakosodott terület orvosi vagy más néven egészségügyi informatika néven vált ismertté. A papíralapú betegnyilvántartás évszázados múltra tekint vissza, és még napjainkban is meghatározó szereppel bír a betegadatok nyilvántartásában. Az egészségügy számos ágában megkezdődött az átállás az elektronikus dokumentáció és nyilvántartás használatára, de ez a folyamat kevésbé haladt előre, mint sok más szektorban. Az egészségügyben az elektronikus dokumentációt főleg adminisztratív célból alkalmazzák, s a döntéshozatalban ezidáig nem jutott lényeges szerephez.
Az elektronikus egészségügyi nyilvántartás (Electronic medical record systems, EMRS) kialakítása az egészségügyi fejlesztések egyik meghatározó területévé vált. Alapvető célja a papírmentes, minden orvosi szakterületet felölelő, minden egészségügyi szolgáltatót lefedő rendszer kialakítása. Ennek egyik fő mozgatórugója, hogy a korábbi rendszer már nem költséghatékony, és a betegellátás színvonalának növelését akadályozza.
Az EMR-rendszerekhez kapcsolódóan többféle, egymást részben átfedő fogalom vált ismertté: az electronic patient record (EPR), az electronic health record (EHR), az electronic medical record (EMR) és a computer-based patient record (CPR). Az EPR elsősorban egy kezelés vagy kivizsgálás eredményeinek a rögzítésére szolgál, míg az EHR akár a beteg teljes kórtörténetét tartalmazhatja. Ennek megfelelően a magyar rendszerben az EPR az ambuláns vagy vizsgálati lapot, illetve zárójelentést válthatja fel, míg az EHR a betegkartont. (Az EPR és az EHR között jelentős az eltérés, mégis gyakran egymás szinonimájaként alkalmazzák ezeket az elnevezéseket. Az EMR a legáltalánosabb jelentésű, mégis leginkább EHR-értelemben használatos.)
Az IOM (Institute of Medicine, USA) 2003-as jelentésében nyolc alapvető funkcionalitást jelölt meg, melyet a kialakítandó EHR-rendszereknek célszerű megvalósítaniuk [12].

- *Kezelési információk és betegadatok.* Közvetlen hozzáférést biztosít a legfontosabb információkhoz a páciens egészségi állapotával kapcsolatban, például korábbi kórképek, diagnózisok, laboratóriumi vizsgálati eredmények, gyógyszerezés, is-

mert allergiák. Mindez javítja az ellátók azon képességét, hogy megfelelő időben helyes klinikai döntéseket hozzanak.

- *Eredménymenedzsment.* Az összes szolgáltató, amely részt vesz a beteg ellátásában, képes az új és a korábbi vizsgálati eredményeket elérni, ezáltal növeli az ellátás hatékonyságát.

- *Rendelésmenedzsment.* Lehetővé teszi gyógyszerreceptek, vizsgálatok és egyéb szolgáltatások megrendelését, illetve tárolását. Mindez elősegíti a szolgáltatáshoz való gyorsabb hozzájutást és a hatékony jogosultság-ellenőrzést.

- *Döntéstámogatás.* Elősegíti a gyógyszerkölcsönhatások feltárását, a diagnózis kialakítását és a megfelelő kezelések megválasztását. Használatával a rendszeres szűrővizsgálatok ütemezése, a betegségmegelőzés hatékonyabbá tehető.

- *Elektronikus kommunikáció és kapcsolattartás.* A hatékony, biztonságos és könnyen hozzáférhető kommunikáció az egészségügyi szolgáltatók között, illetve a betegek és a szolgáltatók között elősegíti az ellátás folyamatosságát, csökkenti a várakozási időt, ezáltal javítja a szolgáltatás minőségét, valamint csökkenti a nemkívánatos események gyakoriságát.

- *A beteg támogatása.* A betegek számára hozzáférést biztosít az őket érintő minden egészségügyi dokumentációhoz, s a betegségekhez kapcsolódóan interaktív tájékoztatást nyújt. Krónikus betegségek esetén segít az otthoni monitoring és önellenőrzés gyakorlatának elsajátításában és fenntartásában.

- *Ütemezési és adminisztratív folyamatok.* Számítógépes felügyeleti eszközök segítségével menetrendek, beosztások készítése, melyek az ellátóhelyek hatékony működését segítik elő.

- *Jelentéskészítés.* Ennek alapja az elektronikus adattárolás, amely egységes szabványokat alkalmazva lehetővé teszi az egészségügyi szervezetek számára, hogy gyorsabban reagáljanak a különféle jelentéstételi kötelezettségekre.

Bár már napjainkban is léteznek rendszerek, melyek e funkciók egy részét megvalósítják, ezek többnyire egymástól szeparáltan működnek, eltérő belső működési folyamatokkal. A hosszú távú cél az EHR-rendszerek fejlesztése során többrétű, az egészségügyi szolgáltatások összehangolásán, hatékonyabbá tételén felül számos más területet érint. Ezek közé sorolható a folyamatok auditálásának lehetővé tétele, az alkalmazott kutatás elősegítése, az egészségügyi oktatás támogatása, vagy akár egy epidemiológiai monitoring rendszer kialakítása. Ezzel együtt az olyan gyakorlati kérdések, mint a rendszerek megfelelő biztonsági szintjének kialakítása, a kapcsolódó jogi és pénzügyi háttér létrehozása még megoldásra várnak. Ugyanakkor az EHR-rendszerek által kezelt adatok struktúrájára, hierarchiájára, tárolásának módjára, a kommunikációs csatornákon keresztül alkalmazandó formátumokra már létrejöttek szabványok (ISO18308, ASTM E31.19, CEN 13606, HL7 [37]), ami azért lényeges, mert az orvosi döntéstámogató rendszereknek EHR-keretrendszerhez megvalósítandó kapcsolódási pontjait, a kialakítandó adatstruktúrákat, és a kommunikáció folyamatát is rögzítik.

7.2. A mesterséges intelligencia szerepe az orvosi döntéstámogatásban

Már a mesterséges intelligencia koncepció kialakulásának korai időszakában felmerült az orvostudomány, mint potenciális alkalmazási terület. Ennek a fő oka az volt, hogy egy-egy orvosi szakterület művelőjének hatalmas mennyiségű tárgyi tudás mellett gyakran évtizedek alatt felhalmozott gyakorlati tapasztalatra volt szüksége ahhoz, hogy feladatait magas szinten legyen képes ellátni. Másképpen fogalmazva egy orvosi szakértő képzése jelentős erőforrás- és időbefektetést igényel, melyet célszerű támogatni minden lehetséges eszközzel. Az orvosi szakértői rendszerek eredeti célkitűzése részben ez volt: a mesterséges intelligencia megfelelő alkalmazásával támogatni az orvosok tevékenységét, legfőképp a diagnózis felállítását. Az 1980-as években a mesterséges intelligencia orvosi szakértői rendszerekben történő felhasználása iránt nagy érdeklődés alakult ki, mely egyben túlzó elvárásokat is támasztott. Egy akkori meghatározás szerint egy mesterséges intelligencia alapú szakértői programnak képesnek kell lennie betegségmodellek és a páciens különböző vizsgálati paraméterei közötti kapcsolatok elemzése alapján diagnózist felállítania és terápiás javallatokat tennie [13]. Jelenleg inkább a laboratóriumi munkát és a tanulást segítik a szakértői rendszerek. Beépítésük a klinikai gyakorlatba jelentős ellenállásba ütközött az orvostársadalom részéről, többek közt azért, mert nem a megfelelő szerepre tervezték e rendszereket. A „mindent eldöntő" gépi szakértő szerepe nem volt elfogadható sem az orvosok, sem a betegek számára. Annak ellenére, hogy a megfelelően finomhangolt szakértői rendszerek nem voltak jelentősen rosszabbak egy átlagos szakértőnél, nem volt meg a kellő bizalom irántuk. Emiatt a kezdeti lelkesedést szkeptikus időszak követte, melynek folyományaként az átfogó szakértői rendszerek helyett később főleg olyan rendszereket készítettek, melyek egy-egy jól meghatározott részfeladatot láttak el, mint például vérösszetétel vizsgálatnál a határértéken túli elemek jelzését.

7.2.1. Tudás alapú következtető rendszerek

Ekkor jelent meg a tudás alapú következtető rendszerek számos formája, melyek a szakértői rendszerek jelentős csoportját képzik. A tudás alapú következtető rendszerek jellemzően a páciens bizonyos paraméterei alapján végeznek egyszerű következtetéseket. Az orvosi tudást legtöbbször szabályok formájában tárolják. Felhasználásuk igen szerteágazó, az alábbi fő területeken kaptak szerepet:

- *Felügyeleti rendszerek:* a beteg állapotát felügyelő műszerekbe integrált szoftver, mely riasztást ad adott bemeneti értékek esetén (például beteg állapotromlásának jelzése).
- *Orvosi képfeldolgozás:* a beteg röntgen-, CT- vagy MRI-felvételeinek kiértékelése alapszinten.
- *Klinikai laboreredmények kiértékelése:* sejttenyészet és szövettani kiértékelés, vér és vizelet-összetevők elemzése.
- *Döntéstámogatás diagnózis felállításához:* a beteg bevitt paraméterei alapján listát ad a lehetséges diagnózisokról (például egészségügyi oktatás támogatása).

- *Terápia-ellenőrzés és -tervezés:* a beteg vizsgálati eredményei és az alkalmazandó kezelési protokollok alapján támogatja a kezelési terv összeállítását, vagy a már elkészített tervet ellenőrzi esetleges új eredmények, tünetek alapján.

Összességében elmondható, hogy a szakértői rendszerek azokon a területeken lettek sikeresek, és váltak mindennapos használati eszközzé, ahol nem voltak szem előtt, vagyis beágyazott módon voltak jelen egy-egy műszerben, mérőeszközben. Ezekben az esetekben csak az adott területspecifikus tudást alkalmazták relatíve egyszerű módon. Különösen a laboratóriumi szakértő rendszerek váltak elterjedtté, mint például a *GermWatcher*, amely sejttenyészetek kiértékelését végzi [14], vagy a *Pathology Expert Interpretative Reporting System, PEIRS* [15], amely többek közt a vérösszetétel elemzését támogatta. A diagnózis felállítását, illetve a terápiás terv készítését támogató rendszerek sokkal kevésbé váltak a mindennapos klinikai gyakorlat részévé, s inkább az orvosi, egészségügyi képzés és oktatás terén kaptak szerepet. Erre jó példa a *DXplain* rendszer [16], melyet eredetileg klinikai döntéstámogatásra készítettek. Tudásbázisa eredetileg 4500 tünetet tartalmazott 2000 lehetséges kapcsolódó betegséggel, aminek alapján képes volt a beteg eredményeit alapul véve rangsorolt diagnózis listát készíteni. Jelenleg főleg oktatási célokra alkalmazzák. A közelmúltban megjelent vagy kialakítás alatt álló EHR-rendszerek célja , hogy beágyazott komponensként segítsék elő az ellátás minőségének javítását. Többek közt betölthetnek ellenőrző, felügyelő funkciókat, mint például a *HELP* rendszer [17], amely riasztást ad le, ha egy tervezett kezeléssel szemben kontraindikáció merül fel egy beteg esetén. (A *HELP* voltaképpen egy korai kórházi információs rendszer, melyet döntéstámogató funkciókkal egészítettek ki.) A korábbi alkalmazási területek mellett jelentős szerepet kaphatnak e rendszerek a kutatás terén. Egyfelől eszközként szolgálhatnak a meglévő orvosi tudás alátámasztására, modellek verifikációjára, másfelől új összefüggések feltárása is lehetővé válik az általuk kezelt nagy mennyiségű információ elemzésével.

7.2.2. Gépi tanulás

A gépi tanulás a mesterséges intelligencia másik jelentős területe, mely az orvosi döntéstámogatás fejlődéséhez nagymértékben hozzájárult. A gépi tanulási technikák központi eleme a tanulás, melynek célja a rendelkezésre álló adathalmazok alapján új összefüggések feltárása, a vizsgált elemek közötti kapcsolatok jellemzése. Végső soron a meglévő tudás bővítésének, elmélyítésének (vagy akár felülvizsgálatának) lehetséges eszközéül szolgálnak ezek a módszerek. Az orvosi adatok sokfélesége a legkülönfélébb technikák alkalmazását igénylik a döntési fáktól a valószínűségi hálókon át a neurális hálókig, de ide sorolhatók a különféle klaszterezési és adatbányászati algoritmusok is. Bár az orvostudomány mint kutatási terület jó táptalajt biztosított a különféle gépi tanulási módszerek kifejlesztéséhez, a napi klinikai gyakorlatban még kevésbé jellemző az alkalmazásuk, mint a tudás alapú szakértői rendszereké. Ehelyett főleg a kutatás során használnak különféle gépi tanulást megvalósító rendszereket. Ugyanakkor számos alkalmazási területen került sor a használatukra a tudás alapú rendszerek tudásbázisának készítésekor, például a KARDIO nevű EKG-elemző szoftvernél [18] a különböző klinikai állapotokat definiáló paraméterek tanulásakor.

7.2.3. Orvosi döntéstámogató rendszerek

Ahogy az az eddigiekből látható, orvosi döntéstámogató rendszerek sokféle formában léteznek, egyrészt tudás alapú következtető rendszerekként, másrészt gépi tanulást megvalósító rendszerekként. Kezdetben önálló rendszereknek szánták őket, és alapvetően a diagnózis felállítása volt a kitűzött fő funkciójuk. A viszonylagos sikertelenséget követően más kórházi rendszerekkel integráltan vagy műszerekbe beágyazottan került sor az alkalmazásukra. Végül Wyatt és Spiegelhalter nyomán a következő definíció vált elfogadottá:
Orvosi döntéstámogató rendszeren (Clinical Decision Support System, CDSS) olyan aktív tudáskezelő rendszereket értünk, melyek képesek két vagy több páciensshez kötődő adat alapján eset-specifikus javaslatokat tenni. [19]
Tehát a korábbi szakértői rendszerek közül a komplex információs rendszerekbe integrált komponenseket tekintették ettől fogva orvosi döntéstámogató rendszereknek. A korábban említett példákon kívül számos más rendszer jelent meg, és került sor alkalmazására [37].
Később a kórházi információs rendszerekkel, illetve azt követően az EHR-rendszerekkel való integrációs törekvést figyelembe véve négy alapfunkciót neveztek meg az integrált orvosi döntéstámogató rendszerek számára [20]:

- *Adminisztratív funkciók:* a kórházi dokumentáció és betegségkódolás támogatása, jogosultságkezelés.
- *Folyamatmenedzsment:* a megfelelő kezelési és kutatási protokollok betartása, elrendelt kezelések nyomon követése.
- *Költségellenőrzés:* gyógyszerezés költségeinek vizsgálata, elrendelt vizsgálatok költség-haszon elemzése.
- *Döntéstámogatás:* diagnózis és kezelési terv kialakításának támogatása, kezelési irányvonalak definiálása.

A döntéstámogatáson felül tehát három másik fő funkció jelent meg. Ezen közül kiemelt hangsúlyt kapott a költségellenőrzés, főként a költség-haszon vizsgálat. Ehhez kapcsolódóan mind a döntéselméleti, mind a hasznosságelméleti aspektusok relevánssá váltak az orvosi döntéstámogató rendszerek kialakítása során, ami egy újabb kapcsolódási pontot jelentett a mesterséges intelligenciához.

7.2.4. Személyre szabott gyógyászat

Több szakértő egybehangzó véleménye szerint a személyre szabott gyógyászat alapját a megfelelően kialakított EMR-rendszerek hálózata jelentheti. Ennek oka, hogy az EMR-rendszerekben tárolt nagy mennyiségű elektronikus egészségügyi-orvosi információ jelentősen elősegítené az élettudományok terén végzett kutatásokat. Lehetővé tenné adott alpopulációra irányuló klinikai vizsgálatok kivitelezését a páciensek előszűrésétől kezdve a hosszú távú nyomon követésig. Az egységes hálózatba épített EMR révén a ritka betegségek esetében is hatékonyabban lehetne megfelelő számú pácienst elérni például egy gyógyszervizsgálathoz. Az eddig többnyire elosztottan rendelkezésre álló

információ egy-egy betegről (családi kórtörténet, betegségek, vizsgálatok eredményei, laboreredmények, beavatkozások, kezelések következményei, hatásossága), illetve betegek tömegeiről az EMR-eken keresztül együttesen elérhető. Ennek magától értetődő kiegészítése lehetne egy ehhez kapcsolódó biobank hálózat, amely a különféle szövet- és vérmintákat tárolná. Különösen daganatos megbetegedések esetén alapvető fontosságú a daganatból származó szövetminta tárolása többek közt a további kutatások elősegítése végett. A kutatási célú felhasználásra példa egy újonnan azonosított genetikai marker tesztelése, vagy egy új kezelési mód vizsgálatához adott klinikai és molekuláris profilú páciensek kiválasztása [35]. Azonban ahhoz, hogy mindez rutinszerűen működhessen, számos fennálló akadályt le kell küzdeni az EMR-nek [33]:

1. *Strukturálatlan, szabad szöveges adatok jelenléte.* Szabványosított adatstruktúrák nélkül a tárolt adatok feldolgozása nehézkes és nem hatékony. Hátráltatja az adatokat felhasználó döntéstámogató eszközöket, illetve az adatok kutatási célú felhasználását.

2. *Eltérő formátumok, különböző belső működés az egyes EMR-eknél.* Előre definiált interfészek és egységesen meghatározott folyamatok kialakítása szükséges az EMR-ek közötti kommunikációhoz. A szabványok egy része már létezik, ezek kötelező használata alapvető feltétele lenne egy EMR-hálózat kialakításának. Az egységesítés nélkül éppen az EMR egyik legnagyobb előnye, a globális (vagy legalábbis a jelentősebb régiókat lefedő) elérhetőség szenvedne csorbát, emiatt nem történne meg a minőségi ugrás az elosztottan, lokálisan rendelkezésre álló adatbázisokhoz képest.

3. *Az egészségbiztosítás rendszerétől függően megszűnhet egy páciens adataihoz való hozzáférés.* Ez a probléma főleg többszereplős egészségbiztosítási piac esetén jöhet létre, ha nem megfelelő a jogi háttér. Ugyanis ha nincs kényszerítve minden egészségbiztosító az egészségügyi-orvosi adatok egymással való megosztására, akkor ennek hiányában egy biztosítóváltás esetén a páciens vagy kezelőorvosa elveszítheti a hozzáférést a korábbi adatokhoz. Mindez a kutatást is hátrányosan befolyásolhatja, például hosszú távú utókövetést végző vizsgálatoknál.

4. *Változó adatminőség.* Bár léteznek standardok, melyek tartalmaznak minőségi követelményeket, még mindig ahány szolgáltató, annyiféle értelmezés és megvalósítás létezik.

Összességében tehát (I.) létre kell hozni és alkalmazni kell az adat standardokat, melyek definiálják az egyes elemek jelentését, azok elvárt struktúráját, illetve lehetséges értékkészletét, (II.) ki kell dolgozni az adatkezelési és adatbeviteli protokollokat, és az ezeknek megfelelő üzleti folyamatokat, végül pedig (III.) a minőséget folyamatosan fenn kell tartani és ellenőrizni kell [35].
Mindeddig a jelenleg megvalósíthatónak ítélt elképzelésekről esett szó, melyek megvalósítására az első próbálkozások már megtörténtek, az első prototípusok már léteznek. Ezeken kívül számos távlati koncepció létezik, melyek közös törekvése, hogy végső soron a betegre bízza, hogy kellő tájékoztatás mellett döntsön, milyen kockázatú és várható hatékonyságú kezelést választ [34]. Ez a megközelítés tehát a pácienst helyezi a középpontba, mint a saját egészségét tudatosan kézben tartani kívánó entitást, aki az általa választott ellátásért fizet.

Mindez természetesen számos technikai, etikai és jogi kérdést vet fel, melyek összességében gyökeres szemléletváltást igényelnek minden érintett (páciensek, egészségügyi dolgozók, biztosítók, szolgáltatók, gyógyszer- és diagnosztikai eszköz-gyártók, szabályozó szervek) részéről. Például az orvosnak egy veszélyesebb vizsgálat előtt megfelelően tájékoztatnia kellene a beteget, hogy az adott teszt milyen kockázattal jár és az eredmény milyen mértékben prediktív. A betegnek pedig képesnek kellene lennie mérlegelnie a kockázat és a prediktív erő szempontjából a vizsgálatot, ami jelenleg talán túlzó elvárás. Ugyanakkor például egy gyógyszer káros mellékhatásainak célzottabb nyomon követése indokolt lenne, azaz a gyógyszer elhagyásán kívül további vizsgálatokat kellene végezni a gyógyszer metabolizmusáról, kiderítendő azt, hogy miért nem volt megfelelő. Egy kimutatás szerint [34] számos gyógyszert vontak ki a forgalomból amiatt, mert a vizsgált teljes populációban nem volt hatékony vagy káros mellékhatásai voltak, azonban létezett olyan alpopuláció, amelyben kiválóan működött. Ha meg lehet ragadni azokat a biomarkereket, amelyek egy-egy gyógyszer hatékony alkalmazását valószínűsítik, akkor a páciensek előszűrésével személyre szabott kezelés válik lehetővé. A tudatos páciens koncepció szerint végeredményben ebben az esetben is a páciensnek kellene döntenie, hogy vállal-e egy olyan kezelést, amely számára nem ideális, illetve mérlegelni azt, hogy a mellékhatásokból eredő kár mikor lépi túl a kezelés nyújtotta előnyöket. Természetesen ennek megvalósításához nagy mértékben változnia kellene a szabályozási környezetnek (ne csak a „mindenkinek jó” gyógyszerek kapjanak engedélyt, hanem olyanok is, melyek alkalmazhatósága egy adott profilhoz kötött), és a gyógyszer- és diagnosztikai eszköz-gyártók hozzáállásának (vállalják el egy-egy gyógyszer alpopulációra szabásának kutatási-fejlesztési költségeit a gyógyszer kidobása helyett).

Egy másik megközelítés szerint az EMR-ek nyújtotta elsődleges előny az, hogy egyes betegségek esetén egyedi kezelési útvonalak kialakítását teszi lehetővé [35]. Itt a páciensek bevonása mellett nagy hangsúly helyeződik a konszenzusos kezelési útvonalak létre hozására, tehát az elérhető nagyszámú eset alapján akkumulálódó közös tudás felhasználására. Egy konszenzusos kezelési útvonal vizsgálatok, kezelések sorrendjét adja meg, ahol az útvonal egyes szakaszait döntési csomópontok határolják el egymástól, melyekből elágazások alakíthatóak ki. Az egyedi kezelés ezeket a vázakat alkalmazza sablonként, és a páciens egyedi jellemzőit felhasználva teszi személyre szabottá. Jelenleg az egyik legnagyobb probléma, hogy a kezelési döntéseknél, például gyógyszerválasztásnál a legtöbbször nem veszik figyelembe a páciens egyedi jellemzőit, hanem ehelyett próba-hiba alapon történik a választás. Még ha egy-egy orvosnál az évek során össze is gyűlik az a fenotípusos információhalmaz, amely orientálhatja a döntést, ezek formális specifikációja rendszerint elmarad. A kezelési útvonalak kialakítása révén viszont lehetővé válna egy adott kezelés alkalmazásakor a sikeres kezelést valószínűsítő fenotípusos vagy genotípusos jegyek meghatározása. Például a mellrák kezelésében már napjainkban is elérhetőek genetikai tesztek, melyek elősegítik a megfelelő terápia kiválasztását, de ezek jelenleg az esetek kis százalékában nyújtanak segítséget. A jövőben feltárt biomarkerek a meglévő kezelési útvonalakat gazdagíthatják, esetleg teljesen át is alakíthatják azokat. Az idő előrehaladtával, ahogy a kezelési tapasztalok gyűlnek, egyre több információ áll majd rendelkezésre ahhoz, hogy a lehető legjobban lehessen személyre szabott terápiát biztosítani az egyes pácienseknek.

A következő alfejezetekben az orvosi döntéstámogató rendszerek alapvető koncepcióihoz kapcsolódó fogalmakat és módszereket mutatjuk be. Elsőként a bináris döntésekhez

kötődő mércéket, majd ezt követően a hasznosságelmélet alapjait tekintjük át, végül pedig a döntési hálókat ismertetjük.

7.3. Bináris döntések kiértékelése

A bináris döntés a legalapvetőbb döntési típus, értelemszerűen két kimenetele lehetséges. A statisztika és gépi tanulás terén szokásos megfogalmazásban ezt a problémát bináris osztályozásnak nevezzük. Ha adott az objektumok egy halmaza, akkor az osztályozó valamilyen algoritmus alapján minden egyes objektumhoz hozzárendel egyet a két lehetséges címke közül, azaz valamelyik osztályba besorolja az objektumot. Abban az esetben, ha rendelkezésre áll a valós címkeinformáció, azaz hogy valójában melyik osztályba tartozik egy-egy objektum, akkor ennek segítségével minősíthetjük az osztályozót. Szokás szerint az egyik osztályt pozitívnak (P), a másikat negatívnak (N) nevezzük. Ez orvosi környezetben egy betegséghez kötődő diagnosztikai osztályozás esetében könnyen értelmezhető: általában a pozitív osztályba sorolt páciensek a betegek, a negatívok a nem betegek. Az osztályozó adott bemeneti paraméter(ek) alapján döntést hoz, ezt nevezzük jósolt értéknek vagy predikciónak. Ez a címke vagy egyezik a valós címkével vagy nem, s ennek megfelelően négyféle kimenetel lehetséges:

1. TP (true positive): Predikció: P, valós érték: P. Az osztályozó helyesen osztályozta pozitívnak a valójában is pozitív besorolású egyedet.

2. FP (false positive): Predikció: P, valós érték: N. Az osztályozó tévesen pozitívnak osztályozta a valójában negatív besorolású egyedet.

3. TN (true negative): Predikció: N, valós érték: N. Az osztályozó helyesen negatívnak osztályozta a valójában is negatív besorolású egyedet.

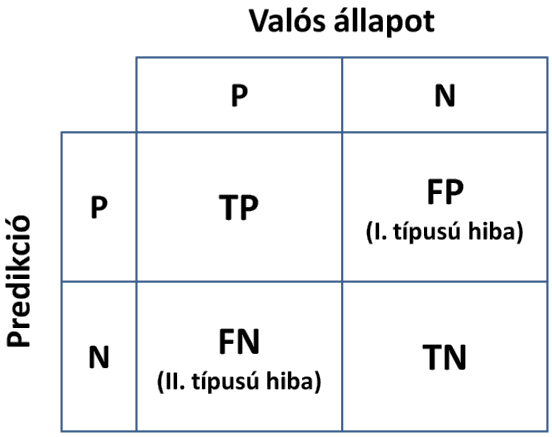
4. FN (false negative): Predikció: N, valós érték: P. Az osztályozó tévesen negatívnak osztályozta a valójában pozitív besorolású egyedet.

E mérőszámokból kiindulva további mércéket alkothatunk, melyek segítségével meghatározhatjuk az osztályozás jóságát.

- Pontosság (accuracy - ACC): a helyes találatok aránya az összes elemhez képest. $ACC = (TP + TN)/(P + N)$
- Érzékenység, szenzitivitás (sensitivity, SENS vagy true positive rate): a helyesen pozitívnak ítélt elemek aránya az összes valójában pozitív elemhez képest. $SENS = TP/(TP + FN)$
- Specifikusság (specificity, SPEC vagy true negative rate): a helyesen negatívnak ítélt elemek aránya az összes valójában negatív elemhez képest. $SPEC = TN/(FP + TN)$
- Hamis pozitívok aránya (false positive rate, FPR): a tévesen pozitívnak ítélt elemek aránya az összes valójában negatív elemhez képest. $FPR = FP/(FP + TN)$

- Pozitív prediktív érték (positive predictive value, PPV): a helyesen pozitívnak ítélt elemek aránya az összes pozitív predikcióhoz képest. $PPV = TP/(TP+FP)$
- Negatív prediktív érték (negative predictive value, NPV): a helyesen negatívnak ítélt elemek aránya az összes negatív predikcióhoz képest. $NPV = TN/(TN + FN)$
- Hamis találatok aránya (false discovery rate, FDR): a tévesen pozitívnak ítélt elemek aránya az összes pozitív predikcióhoz képest. $PPV = FP/(TP + FP)$

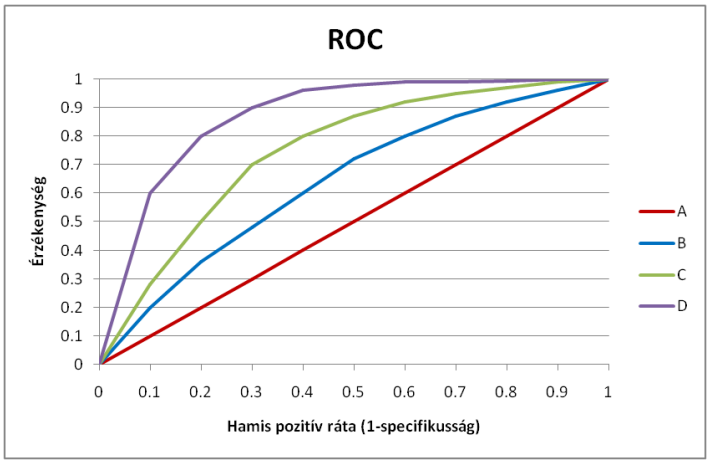
A lehetséges kimenetelek együttesét az úgynevezett *konfúziós mátrix* segítségével is lehet ábrázolni (7.1. ábra). Ebből közvetlenül leolvasható a kétfajta hiba, és ezáltal az osztályozó teljesítményének két fő paramétere. A hamis pozitívok arányát nevezzük I. típusú hibának, a hamis negatívok arányát nevezzük II. típusú hibának. A törekvés a kétféle hiba leszorítására többnyire egymással ellentétes követelményeket támaszt az osztályozóval szemben. Emiatt tipikusan az egyik típusú hibát minimalizálják az osztályozó kialakítása során. Például egy szűréstípusú vírus teszt esetében súlyosabb hiba, ha valakit tévesen egészségesnek, azaz negatívnak jelöl, mint ha valakit tévesen betegnek, azaz pozitívnak jelöl. Az utóbbi esetben ugyanis biztosan további vizsgálatokra sor kerül, tehát fény derülhet a hibára, míg az előbbi esetben esetleg túl későn derül ki az igazság. Tehát ebben az esetben a cél a hamis negatívok arányának leszorítása. A fordított esetre példa egy speciális kezelés alkalmazhatóságát vizsgáló teszt, amely jelentős terhelésnek teszi ki a szervezetet. Ekkor a cél a hamis pozitív esetek leszorítása, mivel ilyenkor azt kell elkerülni, hogy olyan páciens kapjon kezelést, akinél az várhatóan nem használ vagy a szervezete nem viseli el.



7.1. ábra. Konfúziós mátrix

Az osztályozók teljesítményének átfogó minősítésére szolgál a ROC- (*receiver operating characteristic*) görbe, és annak származtatott mennyiségei. A ROC a döntési küszöb változásának függvényében ábrázolja az osztályozó teljesítményét. Az x-tengelyen a hamis pozitív ráta (FPR= 1−SPEC) , az y-tengelyen az érzékenység szerepel. A véletlenszerű osztályozó ROC-görbéje a 45^o-os átlónak feletethető meg (*no discrimination line*). A különböző osztályozók teljesítményét összevethetjük ROC-görbéik segítségével (7.2. ábra). Vizuálisan ez azt jelenti, hogy minél távolabb van egy ROC-görbe a

véletlen osztályozót jelölő átlótól (a felső háromszögben), annál jobb az adott osztályozó teljesítménye. Ha a ROC-görbe az átló alatt fut, az azt jelenti, hogy az osztályozó teljesítménye rosszabb, mint a véletlen. Ebben az esetben viszont invertálni kell a kimenetet, és ekkor egy véletlennél jobb osztályozóhoz jutunk. A ROC-görbe lényege, hogy láthatóvá teszi az igaz pozitívak aránya és a hamis pozitívak aránya közötti egyensúlyozást. Egy átlagos osztályozó esetén, ha úgy állítjuk be a döntési küszöböt, hogy a hamis pozitívak aránya minél alacsonyabb legyen, akkor várhatóan a hamis negatívak aránya nő meg, és vice versa.



7.2. ábra. ROC-görbék. Az A-D görbék egy-egy osztályozó teljesítményét mutatják. Az átlótól legtávolabb lévő D osztályozó nyújtja a legjobb teljesítményt

Az AUC (Area under the ROC curve), azaz a ROC-görbe alatti terület az osztályozó teljesítményének gyakran alkalmazott mérőszáma. Értelemszerűen minél messzebb van (az y-tengelyen felfelé) a ROC-görbe az átlótól, annál magasabb lesz a görbe alatti terület. Előnye, hogy egypontos mérőszámként egyszerűen összehasonlíthatóvá teszi a különféle ROC-görbével rendelkező osztályozókat. Hátránya, hogy nem veszi figyelembe az érzékenység és a hamis pozitívak aránya közötti arányokat a különböző működési tartományokban. Így előfordulhat, hogy egy, a céltartományban jól, máshol viszont gyengén teljesítő osztályozó összességében rosszabb megítélés alá esik az AUC alapján, mint egy, a céltartományban rosszul, máshol átlagosan teljesítő osztályozó. Normált skála esetén az AUC annak a valószínűsége, hogy az osztályozó egy véletlenszerűen kiválasztott pozitív esetet magasabbra rangsorol egy véletlenszerűen kiválasztott negatív esetnél (abban az esetben, ha az osztályozó a pozitív esetekre rendre magasabb pontszámot vagy egyedi értéket ad, mint a negatív esetekre). Az AUC kiszámítható a Gini-index segítségével is az alábbi összefüggések alapján:

$$G = 2 \cdot AUC - 1, ahol \quad G = 1 - \sum_{k=1}^{n} (X_k - X_{k-1}) \cdot (Y_k + Y_{k-1}). \tag{7.1}$$

7.4. Hasznosságelmélet

Számos valós probléma esetén azonban nem lehet bináris döntéseket hozni. Egyfelől a lehetséges döntési opciók száma igen nagy lehet, másfelől gyakran többféle szempont együttesét kell figyelembe venni a kiértékelés során. Az állapotok közötti preferencia megadásához ekkor numerikus mércére van szükség. Ezt a célt szolgálja a *hasznosságfüggvény* (*utility function*) $U(.)$, amely az egyes állapotok kívánatosságának kifejezésére minden állapothoz egyetlen számot rendel. A döntés voltaképpen lehetséges *cselekvések* (*act*) közötti választás. Az egyes cselekvések *kimenetele* (*result*), azaz eredményállapota kívánatosságának megadását teszi lehetővé a hasznosságfüggvény. Egy A cselekvés eredményállapotát $R(A)$, annak hasznosságát $U(R(A))$ jelöli. Nemdeterminisztikus cselekedeteknél többféle eredményállapot állhat elő: $R_i(A)$, ekkor ezek hasznossága: $U(R_i(A))$ mellett a bekövetkezésük valószínűségét: $P(R_i(A))$ is figyelembe kell venni. Az adott cselekvéshez tartozó eredményállapotok valószínűsége és hasznossága szorzatainak összességeként áll elő a cselekvés *várható hasznossága* (*expected utility*). Formálisan a világot leíró tények: E (evidence) ismeretében egy A cselekedet várható hasznossága: $EU(A|E)$ a következő:

$$EU(A|E) = \sum_i P(R_i(A)|A, E) \cdot U(R_i(A)), \tag{7.2}$$

ahol $P(R_i(A)|A, E)$ az i-edik lehetséges eredményállapot bekövetkezésének feltételes valószínűsége A cselekvés végrehajtása esetén, E evidenciák mellett. Racionális megközelítést alapul véve amellett a cselekvés mellett kell döntenünk, amelynek a várható hasznossága a legnagyobb. Ezt nevezzük a *maximális várható hasznosság* elvének (maximum expected utility). Döntéselméleti szempontból ez alapvető fontosságú, mivel segítségével tetszőleges helyzetben meghatározható a megfelelő cselekvés, ha a számítások kivitelezhetőek. Gyakorlati szempontból ugyanakkor a várható hasznosság számítása számos követelményt támaszt. Az evidenciák meghatározásához a világ állapotát valamilyen módon érzékelni kell, az eredményállapotok bekövetkezési valószínűségének számításához a világ állapotai közötti okozati függések ismerete szükséges, egy-egy állapot hasznosságának számítása pedig további komplex összefüggéseken alapulhat. Ezzel együtt a maximális várható hasznosság elve jól használható keretet ad számos döntési probléma megoldására. A hasznosságfüggvények tulajdonságainak leírásához először a racionális preferenciákra vonatkozó megkötéseket kell megismernünk. Tételezzük fel, hogy adott két eredményállapot X és Y. Ekkor e két állapotra vonatkozó preferenciák az alábbiak lehetnek:

- $X \succ Y$: X preferált Y-hoz képest,
- $X \sim Y$: X és Y egyformán preferált,
- $X \succeq Y$: X preferált Y-hoz képest, vagy X és Y egyformán preferált.

Determinisztikus esetben X és Y teljesen specifikált eredményállapotok, míg nemdeterminisztikus esetben egy-egy eloszlást reprezentálnak a lehetséges eredményállapotok halmaza felett. Ez utóbbi esetben X-t és Y-t más néven *szerencsejátéknak* nevezzük. A szerencsejáték a lehetséges kimenetelek: $S_1, S_2, \ldots, S_n$ és azok bekövetkezési valószínűsége: $p_1, p_2, \ldots, p_n$ által alkotott párok halmaza, tehát például $X =$

$[p_1, S_1; p_2, S_2; \ldots, p_n, S_n]$. Ahhoz, hogy a preferenciákra alapozva racionális döntéseket hozhassunk, szükséges szemantikai megkötéseket alkalmaznunk. E megkötéseket más néven a hasznosságelmélet axiómáinak vagy Neumann–Morgenstern axiómáknak nevezzük [21]:

1. Sorrendezhetőség (orderability, completeness). Tetszőleges két állapot: X, Y esetén felállítható egy preferencia sorrend, azaz vagy preferált az egyik állapot a másikkal szemben, vagy egyformán preferált mindkettő. Ez a preferencia megadási feltétel minden lehetőséget lefed, ezért szokás ezt az axiómát teljességi axiómának is nevezni. $(X \succ Y) \vee (Y \succ X) \vee (X \sim Y)$

2. Tranzitivitás (transitivity). Tetszőleges három állapot: X, Y, Z esetén, ha X preferált Y-nal szemben, és Y Z -vel szemben, akkor X-nek preferáltnak kell lennie Z-vel szemben. $(X \succ Y) \wedge (Y \succ Z) \Rightarrow (X \succ Z)$

3. Folytonosság (continuity). Ha adott három állapot: X, Y, Z, ahol Y a preferenciák szempontjából X és Z között helyezkedik el, akkor létezik egy p valószínűség, amely esetén egy racionális döntéshozó számára közömbös, hogy az eredmény egy biztos állapot Y, vagy egy olyan szerencsejáték, amelyben p valószínűséggel X, $1 - p$ valószínűséggel Z az eredmény. $X \succ Y \succ Z \Rightarrow \exists p[p, X; 1 - p, Z] \sim Y$

4. Függetlenség (independence). Adott két szerencsejáték: X és Y, melyek közül X preferált Y-nal szemben, és adott egy irreleváns alternatíva: W, $p \in (0, 1]$. Két olyan összetett szerencsejáték közül, amelyek között a különbség az, hogy az egyikben X helyett Y szerepel, a döntéshozó azt preferálja, amelyik X-et tartalmazza. Mindez a p valószínűségektől és más lehetséges kimenetektől (W) függetlenül igaz. $X \succ Y \Rightarrow [p, X; 1 - p, W] \succ [p, Y; 1 - p, W]$

[22] további két kiegészítő axiómát sorol a hasznosságelmélet axiómái közé, illetve a függetlenségi axiómát egy azzal ekvivalens formában definiálja helyettesíthetőségi axióma néven.

- Helyettesíthetőség (substitutability). Ha két szerencsejáték: X és Y egyformán preferált, akkor a döntéshozó két olyan összetett szerencsejátékot is egyformán preferál, amelyek között a különbség a az, hogy az egyikben X helyett Y szerepel, minden más pedig egyezik. Mindez a p valószínűségektől és más lehetséges kimenetektől (Z) függetlenül igaz. $X \sim Y \Rightarrow [p, X; 1 - p, Z] \sim [p, Y; 1 - p, Z]$

- Monotonitás (monotonicity). Adott két szerencsejáték, melyeknek ugyanaz a két kimenetele lehetséges: X és Y. Ha X preferált Y-nal szemben, akkor ez azt jelenti, hogy azt a szerencsejátékot kell preferálnia a döntéshozónak, amelyik nagyobb valószínűséggel eredményezi X-et. $X \succ Y \Rightarrow (p \geq q \Leftrightarrow [p, X; 1-p, Y] \succeq [q, X; 1 - q, Y]$

- Felbonthatóság (decomposability): Egy összetett szerencsejáték egyszerűbb részekre bontható a valószínűségszámítás szabályai szerint. $[p, X; 1 - p, [q, Y; 1 - q, Z]] \sim [p, X; (1 - p)q, Y; (1 - p)(1 - q), Z]$

A hasznosság ezen axiómái alapján származtatható a hasznosságfüggvény. Ennek érdekessége az, hogy az axiómák csak a preferenciákra adnak megkötéseket, magára a hasznosságra nem.

Definíció Hasznosság elv (utility principle): ha a racionális döntéshozó preferenciái megfelelnek a hasznosság axiómáinak, akkor létezik egy, az eredményállapotokon értelmezett U valós értékű függvény, melyre $U(X) > U(Y)$ akkor és csak akkor teljesül, ha X preferált Y-nal szemben, és $U(X) = U(Y)$ akkor és csak akkor teljesül, ha X és Y egyformán preferált, azaz $U(X) > U(Y) \Leftrightarrow X \succ Y$, $U(X) = U(Y) \Leftrightarrow X \sim Y$.

Mindezek alapján egy szerencsejáték hasznossága az egyes eredményállapotok valószínűségeivel szorzott eredményállapot-hasznosságok összege. Ezt más néven a *maximális várható hasznosság elvének* (*maximum expected utility principle*) nevezzük.

$$U([p_1, S_1; p_2, S_2; \ldots; p_n, S_n]) = \sum_i p_i \cdot U(S_i) \tag{7.3}$$

Tehát ha a lehetséges eredményállapotoknak a hasznosságai és a valószínűségei is specifikáltak, akkor az ezeket tartalmazó összetett szerencsejáték hasznossága teljesen meghatározott, továbbá olyan formában áll elő, mint amit a maximális várható hasznosság elve diktál az 7.3 egyenlet szerint. Másképpen fogalmazva, a preferenciákra vonatkozó kényszerek a maximális várható hasznosság elvének megfelelő hasznosságszámítást eredményeznek szerencsejátékok esetén. Mivel tetszőleges nemdeterminisztikus cselekmény kimenetele szerencsejáték, a maximális várható hasznosság elvén alapuló döntés mindig alkalmazható.

7.5. Hasznosságfüggvények

A hasznosságelmélet alapjai a közgazdaságtanból származnak, ennek megfelelően a hasznosság mérésének első eszközeként a pénz szolgált. A modern gazdaságban szinte minden áru és szolgáltatás rendelkezik pénzben mérhető ellenértékkel, amely közvetve kifejezi a javak kívánatosságát, ez pedig valamilyen mértékig jelzi az adott javak hasznosságát. Mindezek alapján egy egyszerű, intuitív hasznosságfüggvény a rendelkezésre álló pénzmennyiség maximalizálását írná elő. Tehát ha feltételezzük, hogy egy s árut el kívánunk adni, és az egyik vevő X, a másik vevő Y összegű pénzt adna érte ugyanolyan feltételek mellett úgy, hogy $X > Y$, akkor X-et választanánk, mivel ez eredményez nagyobb bevételt. Az azonos feltételek mellett mindig a nagyobb mennyiséget előnyben részesítő preferenciát *monoton preferenciának* hívják. Determinisztikus cselekedetek közötti döntés esetén ez az elv fontos támpontot nyújt, de szükséges megvizsgálni a nemdeterminisztikus esetet is, vagyis a szerencsejátékokat. Felmerül a kérdés, hogy minek alapján határozzuk meg a pénz hasznosságát? Pontosabban, egy rendelkezésre álló pénzmennyiség mennyire befolyásolja a jövőben várható pénzmennyiség hasznosságát? Intuitíve nem ugyanakkora a hasznossága 10 eurónak akkor, ha 1 és akkor, ha 10000 euróval rendelkezünk. Először Bernoulli (1783), majd Grayson (1960) állapította meg, hogy a pénz hasznossága közelítőleg a mennyiségének a logaritmusával arányos. Ebből következően szerencsejáték esetén a várható pénznyeremény hasznossága függ a kiindulási pénzmennyiségtől, továbbá a döntésben szerepet játszik a döntéshozó *kockázatviselési attitűdje* is. Tekintsünk egy olyan szerencsejátékot példaként, amelyben dönteni kell egy a) 50 eurós biztos nyeremény és egy b) 100 eurós lehetséges nyeremény

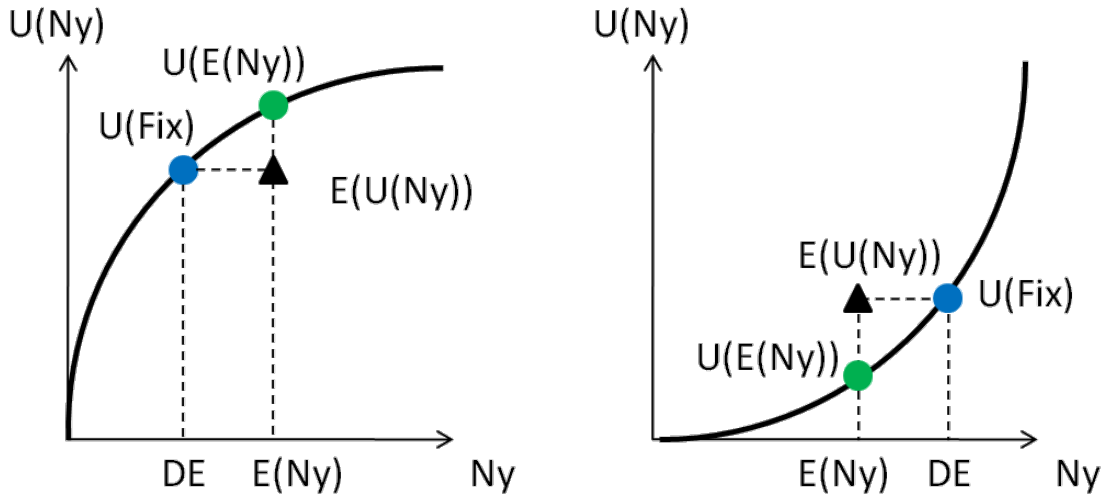
között. Ez utóbbi esetben egy pénzfeldobás kimenetele alapján 100 euró a nyeremény (ha fej) vagy nincs nyeremény (ha írás). A szerencsejáték *várható pénzügyi értéke* (bekövetkezés valószínűsége · pénzmennyiség) azonos a két esetben a) $(1 \cdot 50) = 50$ és b) $(0,5 \cdot 0) + (0,5 \cdot 100) = 50$, a hasznosság megítélésében azonban jelentős különbségek lehetnek a döntéshozók között. Egyesek az alacsonyabb, de biztos nyereményt jobban preferálhatják, míg mások a lehetséges magasabb nyeremény érdekében inkább kockáztatnak. Azt az összeget, amelynek fejében biztosan lemond a döntéshozó a nyereményjátékról, és helyette a biztos nyereményt választja, *determinisztikus ekvivalensnek* nevezzük. A várható pénzügyi érték és a determinisztikus ekvivalens közötti rést *biztosítási (kockázati) prémiumnak* nevezzük. A kockázat kezelése szempontjából három fő típust különböztethetünk meg:

- Kockázatkerülő (risk-averse): a várható pénzügyi értéknél kisebb összegért is eláll a szerencsejátéktól, így a kockázati prémium pozitív. Tehát inkább a biztos nyereményt választja.
- Kockázatsemleges (risk-neutral): nem tesz különbséget a biztos nyeremény és a nyereményjáték között.
- Kockázatkereső (risk-seeking): csak a várható pénzügyi értéknél nagyobb összegért hajlandó elállni a nyereményjátéktól.

Kockázatkerülő magatartás esetén a hasznosságfüggvény konkáv (7.3. ábra), azaz a determinisztikus ekvivalenshez tartozó hasznosság: $U(Fix)$ alacsonyabb, mint a nyereményjáték várható értékének hasznossága: $U(E(Ny))$. Ennek oka az, hogy már „előbb" kiszáll a nyereményjátékból a kisebb, de biztos nyereménnyel, mintsem hogy a nagyobb, de kockázatos nyereményt elérje.
A kockázatkereső magatartás esetén a helyzet pont az ellenkező. A nyereményjáték várható értékének hasznossága: $U(E(Ny))$ alacsonyabb a determinisztikus ekvivalenshez tartozó hasznossághoz: $U(Fix)$ képest. Ez azt mutatja, hogy a kockázatkereső döntéshozó csak akkor fogadja el a biztos nyereményt (determinisztikus ekvivalens), ha annak a hasznossága nagyobb, mint a nyereményjátéké. A nyereményjáték várható hasznossága: $E(U(Ny))$ pedig megegyezik a determinisztikus ekvivalens hasznosságával: $U(Fix)$, ami ez utóbbi definíciójából fakad. Fontos hangsúlyozni, hogy itt a hasznosságról szól az állítás, nem pedig a nyeremény értékéről. A determinisztikus ekvivalenshez tartozó nyereményérték: DE kisebb, mint a nyereményjáték várható értéke: $E(Ny)$ kockázatkerülő magatartás esetén, és $DE > E(Ny)$ kockázatkereső magatartás esetén. Kockázatsemleges magatartás esetén a hasznosságfüggvény közel lineáris: $DE = E(Ny)$. Főleg olyan esetekben jellemző ez, amikor a nyereményjátékon elérhető összeg kis mennyiségű a rendelkezésre álló pénzhez képest.

7.5.1. Hasznosságfüggvények alaptípusai

A hasznosságfüggvényeknek alapvetően két típusát különböztetjük meg: *kardinális* és *ordinális* hasznosságfüggvényeket. A kardinális hasznosságfüggvények numerikus értéket rendelnek minden lehetséges eredményállapothoz. Az állapotok közötti preferenciasorrendet pozitív affin transzformációk (lineáris transzformáció vagy eltolás) mellett is



7.3. ábra. A kockázatkerülő (bal) és a kockázatkereső (jobb) döntéshozó hasznosságfüggvénye. A vízszintes tengelyen látható a nyeremény nagysága NY, a függőleges tengelyen pedig a nyeremény hasznossága $U(NY)$. A nyereményjáték várható értékét $E(NY)$, a várható érték hasznosságát $U(E(NY))$, a nyereményjáték várható hasznosságát pedig $E(U(Ny))$ jelöli. A determinisztikus ekvivalenshez (DE) tartozó hasznosságot $U(Fix)$ reprezentálja

megőrzik. Tehát két hasznosságfüggvény: $U_1(X)$ és $U_2(X)$ között affin transzformáció általi kapcsolat áll fent, ha $U_2(X)$ az alábbi formában áll elő:

$$U_2(X) = a \cdot U_1(X) + b, \tag{7.4}$$

ahol a és b két tetszőleges pozitív konstans, X pedig egy tetszőleges állapot. A transzformáció nem befolyásolja a döntéshozó viselkedését. Az ordinális (vagy sorrendezett) hasznosságfüggvény vagy más néven *értékfüggvény* esetén a numerikus értékek helyett csak a preferenciasorrend van rögzítve. Tehát például két állapot: X és Y esetén $U_1(X,Y)$ megad egy preferenciát X és Y között. Tetszőleges monoton transzformációval (f) előálló $U_2(X,Y)$ megőrzi a preferenciasorrendet, azaz $U_2(X,Y) \equiv f(U_1(X,Y))$. Mindemellett igaz az, hogy ha a döntéshozó preferenciái rögzítettek, akkor a hasznosságaxiómák nem határoznak meg kitüntetett hasznosságfüggvényt.
Számos tárgyterületen a specifikus hasznosságfüggvény mellett gyakran esik a választás általános hibamértékekre, mint a $0-1$ hiba L_0, az abszolút hiba L_1 és a négyzetes hiba L_2. Az ilyen hibacentrikus szemléletű hasznosságfüggvényeket *veszteségfüggvényeknek (loss function, $L(.)$)* nevezzük. Értelemszerűen minél nagyobb a fennálló hiba (veszteség), annál kisebb a hasznosság. Az optimális értékek ekkor rendre a következőek:

$$L_0(x,\hat{x}) = I(x \neq \hat{x}) : módusz \tag{7.5}$$
$$L_1(x,\hat{x}) = |x - \hat{x}| : medián \tag{7.6}$$
$$L_2(x,\hat{x}) = (x - \hat{x})^2 : átlag, \tag{7.7}$$

ahol $\hat{x}$ a referencia érték, x az aktuálisan vizsgált (jelentett) érték és $I(.)$ egy indikátor-

függvény. Ha ezek az értékek értelmezhetőek egy diszkrét valószínűségi eloszlásként $\hat{p}, p$, akkor alkalmazható a keresztentrópia $H(p\|\hat{p})$ és a Kullback–Leibler távolság $KL(p\|\hat{p})$ mint veszteségfüggvény:

$$H(p\|\hat{p}) = -\sum_i p_i log(\hat{p_i}), \tag{7.8}$$

$$KL(p\|\hat{p}) = \sum_i p_i log(p_i/\hat{p_i}). \tag{7.9}$$

A továbbiakban a tárgyterületspecifikus QUALY és a micromort hasznosságfüggvényeket vizsgáljuk meg, melyeket gyakran alkalmaznak orvosi és biztonsági elemzésekben az emberi élet és egészség (hasznosságának) számszerűsítésére.

7.5.2. QUALY

A QUALY azaz quality-adjusted life year [25] az egyik elterjedten alkalmazott hasznosságfüggvény az egészségügyben (például az Egyesült Királyság területén). A QUALY alapvetően két tényezőt vesz alapul: egyfelől hány évvel hosszabbítja meg egy orvosi kezelés a páciens életét, másfelől milyen életminőség-változást tesz lehetővé. Az életminőség meghatározása egy 0-tól 1-ig terjedő skálán történik, ahol 1 jelenti a tökéletes egészségi állapotot, 0 pedig a halált. A nehézséget annak meghatározása okozza, hogy az életminőséget csökkentő egyes tényezők (például állandó fájdalom) milyen mértékű csökkenést eredményezzenek, s a tökéletes egészség meghatározása sem egyértelmű. Bár léteznek standardizált kérdőívek a QUALY meghatározásához (például EQ-5D [24]), az értékek megadásakor nem lehet kizárni a szubjektivitást. Az életminőségváltozás becsléséhez először a jelenlegi életminőséget mérik fel, amelynek során különféle tényezőket vesznek figyelembe, például általános közérzet, fájdalmak jelenléte, mozgáskészség. Ez után számítják ki a kezelést követő vagy annak folyamán végbemenő lehetséges állapotváltozás hatását. Ekkor kerül sor a QUALY-hoz kapcsolódó alapvető fontosságú mutató, a kezelés költségének számítására, amelyet jellemzően 1 egységnyi QUALY-ra vetítenek. A mutatót legtöbbször krónikus vagy terminális betegségek kezelésének eldöntéséhez használják (például daganatos megbetegedések esetén), melyeknél ez a költség számottevő. A QUALY segítségével összehasonlítható az életet meghosszabbító, de az életminőségét nem javító kezelés a pusztán életminőséget javító terápiával. Használható a költségek visszafogására is a nem kellően költséghatékony terápiák elutasításával. Egy általános elfogadási küszöb kiválasztása nem lehetséges, mivel a küszöb számos tényezőtől függ, legfőképpen a rendelkezésre álló erőforrásoktól. A küszöb létezése önmagában is nehezen kezelhető kérdéseket vet fel. A QUALY alkalmazását érő számos kritika egy része a megfelelő küszöbérték megválasztására koncentrálódik, másfelől arra, hogy nem tesz különbséget eltérő súlyosságú betegségek között, ha a páciensek QUALY-változása azonos. Tekintsünk egy hipotetikus példát a QUALY alkalmazására, az egyszerűség kedvéért forintban számolva. Tegyük fel, hogy egy páciensnek súlyos daganatos betegsége van, amely alacsony várható élettartammal jár. Két lehetőség a kezelésére: a hagyományos kemoterápia és egy új gyógyszeres kezelés. Az előbbi egy évre növeli a várható élettartamot, azonban az életminőséget jelentősen rontja (0,4), így összességében a QUALY $1 \cdot 0,4 = 0,4$. Az utóbbi kezelés

másfél évre növeli a várható élettartamot, és a hagyományos terápiához képest jobb életminőséget tesz lehetővé (0,6), így összességében a QUALY $1{,}5 * 0{,}6 = 0{,}9$. A hagyományos kezelés költsége 1 MFt, míg az új kezelésé 6 MFt. A két kezelés közötti eltérés tehát 0,5 QUALY és 5 MFt költség, így egységnyi QUALY-ra 5 MFt/0,5 = 10 MFt esik. Ez azt jelenti, hogy a páciens nagy valószínűséggel nem kapja meg az új gyógyszert (nem finanszírozza az illetékes egészségbiztosító), mivel az nem lenne költséghatékony. Összehasonlításképpen megemlítjük, hogy 2010-ben az Egyesült Királyságban 20 000 és 30 000 font (7,12 – 10,67 MFt) per QUALY között helyezkedett el a költséghatékonysági küszöb [27].

7.5.3. Micromort

A micromort kismértékű kockázatok reprezentálására és összevetésére szolgáló mennyiség, melynek 1 egysége (1 μmrt) egy az egymillióhoz esélyt jelent az elhalálozásra ($p = 10^{-6} = 1\mu mrt$). A mértékegységet, Ronald Howard alkotta meg a 1970-ben [29]. A micromort lehetővé teszi eltérő tevékenységi területekről származó alacsony kockázatú események összehasonlítását, ilyen lehet például egy kórházi szülés (80 μmrt) és 100 km út megtétele motorral (11μmrt) [28]. A micromort számításához használt időegység az alkalmazási területtől függ. Leggyakrabban az egy napra eső kockázatot és az adott tevékenység teljes időtartamára eső összkockázatot szokták számítani. Az Egyesült Királyságból származó 2008-as statisztikai adatok szerint az egy napra eső nem természetes halál kockázata nagyjából 1 micromort (mintegy 18 000 nem természetes halál egy év alatt, 54 000 000 a becsült össznépesség, így $18\,000/(54\,000\,000 \cdot 365) = 0{,}0000009132 \approx 1\mu mrt$)) [28]. Az egészségügyben a két leggyakrabban használt micromort egység a beavatkozásra számított és a kórházi napok számára vetített kockázat. A beavatkozásra számított kockázatra példa a már említett szülés, vagy a műtétet megelőző általános érzéstelenítés és altatás kockázata, ami megközelítőleg 10 micromort. A kórházban töltött időszak önmagában kockázatos lehet egy esetleges betegbiztonságot veszélyeztető nem megfelelő kezelés miatt vagy az odafigyelés hiánya miatt, s kockázati tényező lehet a kórházban szerzett fertőzés is. A statisztikák alapján a kellő odafigyeléssel megelőzhető kórházi elhalálozások kockázata 75 micromort [28]. Hasznosságfüggvényként a micromort akkor használható, ha azt a pénzmennyiséget vesszük alapul, amelyet az ember hajlandó lenne kifizetni a kockázat elkerülése érdekében. A 2009-es adatokat figyelembe véve 1 micromort 50 $-nak feleltethető meg (az USA-ban) [23].

7.6. Többváltozós hasznosságfüggvények

A legtöbb valós problémánál egyszerre több szempontot kell figyelembe venni a döntéshozás során. Tekintsünk példaként egy orvosi döntéshozatalt, amikor is egy balesetben térdsérülést szenvedett páciens kezelési-rehabilitációs tervének a kialakítása a cél. Ekkor figyelembe kell venni, hogy a lehetséges kezelési módok milyen mértékű felépülést valószínűsítenek, mekkora az egészségromlás (vagy esetleg a halál) kockázata a kezelés következtében, továbbá a kezelés költsége is szempont lehet. Az ilyen több attribútummal leírható problémák a *többattribútumos hasznosságelmélet* (multiattribute utility theory) segítségével kezelhetőek. Egy attribútum rendelkezhet diszkrét vagy folytonos

értékkel. Az egyszerű értelmezés érdekében az attribútumok értékkészletét a legtöbbször úgy határozzuk meg, hogy a nagyobb értékhez nagyobb hasznosságérték tartozzon. Általános esetben a hasznosság meghatározása az egyes attribútumkombinációk alapján igen összetett feladat lehet. Vannak azonban speciális esetek, amikor nem szükséges meghatározni a konkrét hasznosságot, mert a hasznosságérték nélkül is dönteni lehet az alternatívák között. Jelölje $\underline{R} = R_1, R_2, \ldots, R_n$ az attribútumokat, X_{R_1} pedig az X állapot R_1 attribútumát. Az egyik ilyen eset, amikor két eredményállapot (X és Y) közül az egyik (X) minden attribútumában kedvezőbb a másiknál, azaz

$$\forall_i : X_{R_i} \succ Y_{R_i} \quad i = 1, \ldots, n. \tag{7.10}$$

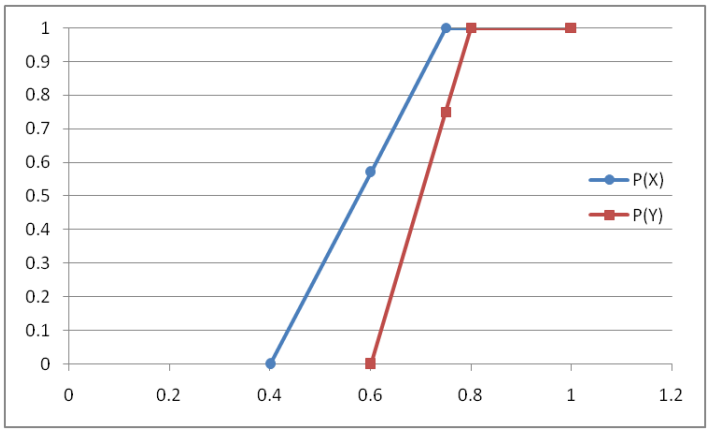
Ezt úgy nevezzük, hogy X *szigorúan dominálja* Y-t. A korábbi térdsérüléses példánál maradva vizsgáljunk meg két lehetséges kezelési módot a költség és a biztonság attribútumok alapján. Tegyük fel, hogy X egy hagyományos gyógyszeres kezelés, míg Y egy újonnan kidolgozott mágnesterápiás eljárás. Ha ez az Y új terápia olcsóbb, mint az X korábbi kezelés, és biztonságosabb is, akkor ez azt jelenti, hogy Y szigorúan dominálja X-et, azaz jelen esetben a hasznosságérték számítása nélkül is képes dönteni a döntéshozó. A szigorú dominancia előnye, hogy leszűkíti azon állapotok számát, melyek között a végső döntést meg kell hozni, azonban nem feltétlenül eredményez egyértelmű döntést. Determinisztikus esetben a szigorú dominancia jól alkalmazható, nemdeterminisztikus esetben azonban nem végezhetünk ilyen egyértelmű módon összehasonlítást, mivel ilyenkor minden alternatíva összes lehetséges kimenetelét figyelembe kellene venni. Ekkor vizsgálhatjuk a *sztochasztikus dominancia* meglétét, ami a dominanciatulajdonság nemdeterminisztikus általánosítása. A sztochasztikus dominancia megállapításához meg kell vizsgálni a szóban forgó nemdeterminisztikus eredményállapotok eloszlásfüggvényét. Ha X és Y nemdeterminisztikus cselekvések kimenetele az R attribútumon értelmezve a $p_X(R)$ és $p_Y(R)$ valószínűségi eloszlásokkal áll elő, akkor Y sztochasztikusan dominálja X-t, ha

$$\forall R \int_{-\infty}^{R} p_Y(r)dr \leq \int_{-\infty}^{R} p_X(r)dr, \tag{7.11}$$

azaz Y eloszlásfüggvénye mindig jobbra esik X eloszlásfüggvényétől. A térdsérülés kezelésének példáját folytatva hasonlítsuk össze a két kezelési alternatívát a várható felépülés mértéke szerint. Tegyük fel, hogy a hagyományos kezelés (X) esetén a felépülés mértéke egyenletes eloszlású, $40 - 75\%$ között várható, míg a mágneses terápia esetén (Y) ugyanez $60 - 80\%$ között várható ugyancsak egyenletes eloszlással. Ekkor teljesül az, hogy Y eloszlásfüggvénye mindig „jobbra található” X-től (lásd 7.4. ábra), vagyis Y sztochasztikusan dominálja X-et.
Mindennek a jelentősége az, hogy ha Y sztochasztikusan dominálja X-et (minden attribútum esetén), akkor tetszőleges $U(.)$ monoton nem csökkenő hasznosságfüggvény esetében $EU(Y) \geq EU(X)$, azaz Y várható hasznossága legalább akkora, mint X-é. Tehát X figyelmen kívül hagyható.

7.6.1. A preferenciák strukturáltsága

A korábban említett általános esethez képest jelentős egyszerűsítést jelenthet a hasznosság számítása során, ha a döntéshozó preferenciái strukturáltságot mutatnak. Ellenkező



7.4. ábra. Sztochasztikus dominancia. Y eloszlásfüggvénye mindig jobbra található X-től, azaz Y sztochasztikusan dominálja X-et

esetben, például n darab attribútum esetén, ha minden egyes attribútum k különböző értéket vehet fel, a teljes $U(r_1, r_2, \ldots, r_n)$ hasznosságfüggvény megadásához a legrosszabb esetben k^n értékre lenne szükség. A preferenciák strukturáltságának vizsgálatakor meg kell különböztetni a determinisztikus és a nemdeterminisztikus környezetet. Determinisztikus esetben, ordinális hasznosságfüggvényeknél jelentős egyszerűsítés érhető el, ha az attribútumok *kölcsönösen preferenciálisan függetlenek*. Jelölje $\langle r_1, r_2, r_3 \rangle$ az R_1, R_2, R_3 attribútumok értékeinek vektorát. Két attribútum, R_1 és R_2 preferenciálisan függetlenek R_3-tól, ha az $\langle r_1, r_2, r_3 \rangle$ és $\langle r_1', r_2', r_3' \rangle$ közötti preferencia nem függ R_3 adott értékétől. Abban esetben, ha minden $R_i, R_j, R_z \in \underline{R} = R_1, R_2, \ldots, R_n, i \neq j \neq z$ attribútumra teljesül, hogy R_i és R_j preferenciálisan független R_z-től, akkor a $\underline{R}$ attribútumok kölcsönösen preferenciálisan függetlenek. Mindennek a jelentősége az, hogy ha az $R_1, R_2, \ldots, R_N$ attribútumok kölcsönösen preferenciálisan függetlenek, akkor a döntéshozó viselkedési preferenciája leírható a következő függvény maximalizálásával:

$$V(r_1, r_2, \ldots, r_n) = \sum_{i=1}^{n} V_i(r_i), \tag{7.12}$$

ahol minden V_i egy értékfüggvényt jelöl, ami az adott R_i függvénye, V pedig az összetett ordinális hasznosságfüggvény, azaz értékfüggvény. Ha feltételezzük, hogy a korábbi példában a kölcsönös preferenciafüggetlenség fennáll, és determinisztikus a környezet, akkor a költség (ezer Ft), a biztonság és a kezelés hossza (napok) attribútumok alapján egy lehetséges ordinális hasznosságfüggvény a következő: *V(Költség, Biztonság, Kezelési idő) = - Költség + Biztonság* $\cdot 10^2$ *- Kezelési idő.* Az ilyen alakban előálló értékfüggvényt *additív értékfüggvénynek* nevezzük, s ezek jól használhatók valós problémáknál is közelítésként. Nemdeterminisztikus környezetben a helyzet összetettebb, mivel ekkor a szerencsejátékok kimenetelét, illetve az ezek közötti preferenciákat, és a kardinális hasznosságfüggvényeket is kezelni kell. Mindezt lehetővé teszik a korábban ismertetett fogalmak kiterjesztései, a *hasznosságfüggetlenség* és a *kölcsönös hasznosságfüggetlenség.*

„Az attribútumok X halmaza hasznosságfüggetlen az attribútumok Y halmazától, ha az X attribútumokon alapuló szerencsejátékok közötti preferenciák függetlenek az Y-beli

attribútumokhoz rendelt értékektől." [22]
Kölcsönös hasznosságfüggetlenség értelemszerűen akkor áll fenn egy adott attribútumhalmazra, ha annak minden részhalmazára fennáll a hasznosságfüggetlenség. A kölcsönös hasznosságfüggetlenség előnye, hogy teljesülése esetén *multiplikatív hasznosságfüggvénnyel* írható le a döntéshozó viselkedése. Két attribútum: R_1 és R_2 esetén például az alábbi formát veszi fel a hasznosságfüggvény:

$$U(R_1, R_2) = s_1 \cdot U_1(R_1) + s_2 \cdot U_2(R_2) + s_1 \cdot s_2 \cdot U_1(R_1) \cdot U_2(R_2), \tag{7.13}$$

ahol s_1, s_2 állandók.
Minden egyattribútumos hasznosságfüggvény a további attribútumoktól függetlenül alakítható ki. Összességében a döntéshozó preferenciarendszere e hasznosságfüggvények kombinációival révén írható le.

7.7. Döntési hálók

A racionális döntések elemzésének gyakran alkalmazott eszközei a döntési hálók vagy más néven hatásdiagramok. A döntési háló lehetővé teszi az egyes cselekvési alternatívákhoz tartozó eredményállapotok valószínűségének és hasznosságának vizsgálatát. Felépítésüket tekintve a valószínűségi hálók kiterjesztéseinek tekinthetők, és alkalmasak a hasznosság kezelésére. Háromféle csomóponttípussal rendelkeznek:

1. *Valószínűségi csomópontok (chance nodes).* A valószínűségi hálókhoz hasonlóan ez a csomóponttípus valószínűségi változókat jelöl (ellipszis formában). Egy-egy változó lehetséges értékeit a szülőcsomópontok értékein alapuló feltételes valószínűség eloszlás adja meg, amelyet leggyakrabban feltételes valószínűségi táblázat formájában tárolnak. Szülőcsomópontja (olyan csomópont, ahonnan irányított él fut bele) lehet döntési csomópont vagy más valószínűségi csomópont. A feladata, hogy reprezentálja egy döntés következményeként kialakuló lehetséges eredményállapotokat.

2. *Hasznosságcsomópontok (utility nodes).* A döntéshozó hasznosságfüggvényét reprezentálják a szülőcsomópontokon definiált függvény segítségével, amely vagy táblázatos formában adható meg vagy egy parametrikus függvényként (ami lehet additív vagy lineáris). Szülőcsomópontjai olyan állapotoknak felelnek meg, melyek befolyással vannak a hasznosság megítélésére. Vizuálisan egy rombusz reprezentálja.

3. *Döntési csomópontok (decision nodes).* E négyszöggel jelölt csomópontok azokat a döntési pontokat reprezentálják, ahol a döntéshozónak választani kell a lehetséges alternatívák közül.

Számos gyakorlati esetben a teljes döntési háló igen komplex lehet, emiatt gyakran egyszerűsített formát, az úgynevezett *cselekvéshasznosság reprezentációt* használunk. Ekkor a kimeneti állapot elhagyásával a hasznosságcsomópont a jelenlegi állapotot leíró döntési és valószínűségi csomópontokhoz kapcsolódik. A hasznosságcsomópont ebben az esetben a döntési alternatívákhoz (cselekvésekhez) kapcsolódó várható hasznosságokat definiálja cselekvéshasznosság táblák segítségével.

7.7.1. Döntési hálók kialakítása és kiértékelése

A döntési háló kialakításának lépései hasonlóak a valószínűségi háló építéséhez, azzal a különbséggel, hogy döntési háló esetében a döntési és a hasznosságcsomópontokat is létre kell hozni, és gondoskodni kell a felparaméterezésükről. Az alábbi fő lépések [22] alkotják ezt a tudásmérnöki folyamatot.

- *Oksági modell létrehozása.* A vizsgált problémához kapcsolódó változók azonosítása az első lépés. Ezt nagymértékben befolyásolja, hogy milyen részletezettségű döntési hálóra van szükség, illetve mely változók elérhetőek vagy megfigyelhetőek. A változóknak megfelelően egy-egy csomópontot veszünk fel. A változók között lévő kapcsolatok meghatározása a következő lépés. Ennek során a közvetlen ok-okozat kapcsolatban lévő változókat reprezentáló csomópontokat irányított élekkel kötjük össze ok $\Rightarrow$ okozat irányban. Mindez alapulhat szakirodalmi adatokon vagy szakértők tudásán.

- *Kvalitatív döntési modell kialakítása.* A problématerületet leíró oksági modell gyakran túl komplex az adott döntési feladat hatékony megvalósításához, ezért ilyenkor egyszerűsítésre van szükség. A döntést közvetlenül nem befolyásoló változók elhagyhatók, mások felbonthatók vagy összevonhatók. Ennél a lépésnél ügyelni kell a majdani alkalmazási területen jellemző döntési mechanizmusokra, azaz például melyek a tipikus kimeneti változók, illetve milyen kvantáltsági szint az elvárt. Mindennek célja az, hogy a háló olyan döntési helyzetben nyújtson segítséget, amely ténylegesen előáll.

- *Valószínűségek meghatározása.* A struktúra rögzítését követően a valószínűségi csomópontok feltételes valószínűség eloszlásának a specifikálása a következő lépés. Ez az esetek többségében feltételes valószínűség táblák kitöltését jelenti. Az egyes (feltételes) valószínűségek szakértők becslései vagy szakirodalmi adatok alapján adhatók meg. Az ok-okozat irányú becslés előnye, hogy az emberi becslést befolyásoló tényezők is ezt torzítják a legkevésbé, szemben a diagnosztikai irányú (következményről kiváltó okra következtető) becsléssel.

- *Hasznosságok rögzítése.* A hasznosságok meghatározásához szükséges a lehetséges eredményállapotok közötti preferencia, amely többek közt szakértői véleményen alapulhat. A numerikus hasznosságértékek hozzárendelése történhet a lehető legjobb és a legrosszabb eredményállapot között felállított skála arányos beosztása szerint, ha a lehetséges eredményállapotok kisszámúak. Komplexebb esetben többattribútumos hasznosságfüggvény megadására is szükség lehet.

- *A modell finomhangolása.* A kialakított modell teljesítményének a méréséhez szükség van valamilyen referenciára (azaz bemenet-kimenet párokra), amely valamiféle elvárt viselkedést tükröz. Erre egy lehetőség a modell alapján adódó döntéseket összevetni egy, a területen jártas szakértő döntéseivel. A cél a nem megfelelő, avagy nem jól hangolt modellelemek azonosítása és kijavítása.

- *Érzékenységvizsgálat.* A modell validálása folyamán külön figyelmet kell fordítani az érzékenységvizsgálatra, melynek lényege, hogy feltárja a valószínűségi értékek,

illetve hasznosságok változtatásának hatását. Ehhez különböző beállítások mellett kell megvizsgálni a kiadódó eredményt. Mindezek alapján megállapítható, hogy az egyes valószínűségek kismértékű változására mennyire érzékeny a legjobb döntés. Ha kis változtatás a „bemeneten" jelentős változást okoz a „kimeneten", akkor lehetséges, hogy nem megfelelő a változó(k) részletezettségi szintje, vagy hiányos a változók közötti kapcsolatok reprezentációja. Viszont ha egy változóhoz vagy változócsoporthoz kapcsolódó valószínűségek nagymértékű változása csak kis mértékben nyilvánul meg, akkor lehet, hogy a változók összevonhatóak, elhagyhatóak, vagy elegendő egy közelítő becslés. Az érzékenységi vizsgálat előnye, hogy jelzi, hol nem megfelelő a valószínűségek numerikus becslése.

Az elkészült döntési háló kiértékelése a döntési csomópontok lehetséges értékei mentén történik, azaz minden lehetséges értékre el kell végeznünk a háló kiértékelését. Ha a döntési csomópontnak értéket adunk, akkor egy rögzített ténybeállítású valószínűségi csomópontként viselkedik. A döntési háló kiértékelését végző algoritmus a következő lépéseket valósítja meg:

- A jelenlegi állapotnak megfelelő evidenciák beállítása a csomópontokon. Tehát az érintett csomópontok felveszik az evidenciáknak megfelelő értéket.
- A döntési csomópont minden egyes értékére:
  - A döntési csomópont rögzítése az adott értéken.
  - A hasznosságcsomópont szüleire a posteriori valószínűségek számítása egy (szabványos) valószínűségi következtető algoritmussal.
  - A döntési csomópont adott értékéhez (egy cselekvéshez) tartozó hasznosság kiszámítása.
- Az eltárolt hasznosságértékek közül visszaadja a legnagyobbat.

7.7.2. Döntési hálók tulajdonságai

A döntési háló előnyei leginkább akkor érzékelhetőek, ha összevetjük egy másik gyakran alkalmazott döntéstámogatási eszközzel, a döntési fával. A döntési fa explicite tartalmazza egy döntési szituáció összes lehetséges kimenetelét, emiatt komplex, sokténye zős problémáknál a fa kezelhetetlenül nagy lesz, ami mind az információ tárolása, mind a kiértékelés során nehézségeket okoz. Ezzel szemben a döntési háló kompakt módon képes reprezentálni a tárgyterületi tudást. Ugyanakkor erősen aszimmetrikus problémák vizsgálatához a döntési fa alkalmasabb lehet, mivel döntési hálók struktúráján az aszimmetria nem mutatkozik meg. Esetenként az események időzítését egyszerűbb nyomon követni egy döntési fás reprezentációval, bár a döntési hálóknál is expliciten megjelenik ez a információ a struktúrában [36].
A döntési hálók kiértékelésére léteznek hatékony algoritmusok. Ilyen a változóeliminálás algoritmus [32], a Shachter-algoritmus [26] és az ezzel ekvivalens élfordítás (arc reversal) algoritmus [31]. Ez utóbbiak gráftranszformációs lépések sorozata révén redukálják a döntési hálót addig, amíg csak egy hasznossági csomópont marad. Ekkor

a csomópont egy adott döntési alternatívához tartozó hasznosságot adja meg. Az algoritmus négy operátort használ, melyek egymást követő (a szabályok sorrendjében történő) alkalmazásával valósul meg a döntési háló redukciója és ezáltal kiértékelése. A négy lehetséges művelet a következő:

1. *Zárvány csomópontok eltávolítása.* Egy valószínűségi csomópont zárványnak tekinthető, ha nincs gyermeke. A zárvány csomópontok egyszerűen eltávolíthatóak a döntési hálóból.

2. *Döntési csomópont eltávolítása.* Egy D döntési csomópont eltávolítható, ha egyetlen gyermeke U hasznosságcsomópont. Ekkor a $D \to U$ él helyett D szüleiből $\underline{Pa_D}$ egy-egy élt veszünk fel U-ba. Jelölje $\underline{Pa_U}$ az U csomópont szüleit, illetve a műveletet megelőző hasznosságot $U_{old}(\underline{Pa_U})$. Az új hasznosság a maximális hasznosságú döntési alternatíva d választásával adódik:

$$U_{new}(\underline{v}) = max_d U_{old}(\underline{pa_U}), \tag{7.14}$$

   ahol $\underline{V} = \underline{Pa_D} \cup \underline{Pa_U} \setminus D$.

3. *Valószínűségi csomópont eltávolítása.* Egy valószínűségi csomópont X akkor távolítható el a hálóból, ha az egyetlen gyermeke az U hasznosságcsomópont. Jelölje $\underline{Pa_X}$ az X csomópont szüleinek halmazát, $\underline{Pa_U}$ pedig az U csomópont szüleinek halmazát, továbbá $P_{old}(x|\underline{pa_X})$ X feltételes valószínűségét a csomópont eltávolítás előtt és $U_{old}(\underline{pa_U})$ a korábbi hasznosságot. X eltávolításkor az X-be futó minden egyes él helyett $\underline{Pa_X}$ minden egyes eleméből U-ba húzunk élt. Az új hasznosság a következőképpen áll elő:

$$U_{new}(\underline{v}) = \sum_x P_{old}(x|\underline{pa_X}) \cdot U_{old}(\underline{pa_U}), \tag{7.15}$$

   ahol $\underline{V} = \underline{Pa_X} \cup \underline{Pa_U} \setminus X$.

4. *Élfordítás.* Egy $X \to Y$ él megfordítható $Y \to X$ éllé, ha nincs más irányított él X és Y között. Ekkor X szülei Y szüleivé válnak és vice versa. Ekkor Y új feltételes valószínűsége:

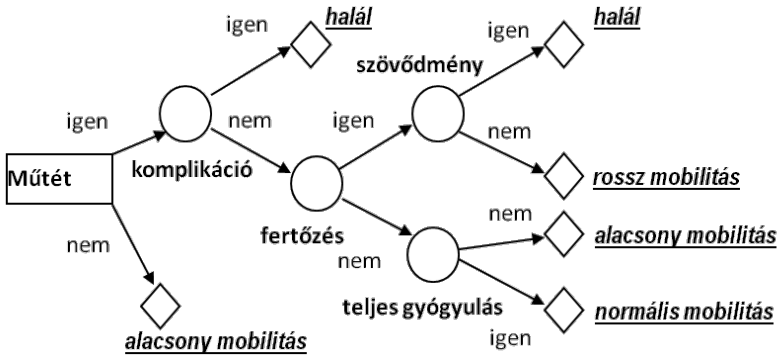
$$P_{new}(y|\underline{v}, \underline{w}, \underline{z}) = P_{old}(x|\underline{v}, \underline{w}) \cdot P_{old}(y|\underline{w}, \underline{z}), \tag{7.16}$$

   X új feltételes valószínűsége pedig:

$$P_{new}(y|x, \underline{v}, \underline{w}, \underline{z}) = \frac{P_{old}(x|\underline{v}, \underline{w}) \cdot P_{old}(y|\underline{w}, \underline{z})}{\sum y P_{new}(y|\underline{v}, \underline{w}, \underline{z})}, \tag{7.17}$$

   ahol $\underline{V} = Pa_X \setminus Pa_Y$, $\underline{Z} = Pa_Y \setminus Pa_X \cup X, \underline{W} = Pa_Y \cup Pa_X$, az $\underline{V}$ kizárólagosan X, $\underline{Z}$ kizárólagosan Y szüleit tartalmazza, míg $\underline{W}$ tartalmazza a közös szülőket.

Tekintsük a korábbiakban már említett térdkezelésről szóló példát. Adott egy páciens, aki súlyos térdbántalmakban szenved, ennek köszönhetően mobilitása (mozgáskészsége) alacsony, kezelőorvosa a műtétet mérlegeli. Ebben az esetben a hasznossági csomópont a mozgáskészség különböző fokozataihoz rendel egy hasznosságértéket (U), a döntési csomópont pedig a műtét elvégzése. A műtét az esetek nagyon kis részénél súlyos komplikációkkal jár ($p = 0{,}01$), melyet a beteg nem él túl ($U(\text{halál}) = 0$). Az esetek egy további hányadánál fertőzés alakulhat ki ($p = 0{,}045$), ami súlyos szövődményekkel járhat. A legrosszabb esetben a beteg nem éli túl ($p = 0{,}07, U = 0$), vagy maradandó károsodást szenved ($p = 0{,}93$), ami korlátozza a mozgáskészségét (U(rossz mobilitás)= 2). A legtöbbször azonban semmilyen komplikáció nem lép fel ($p = 0{,}955$), ekkor a beteg akár teljesen visszanyerheti a normális mozgáskészségét ($p = 0{,}65$, U(normális mobilitás)=10), de az is meglehet, hogy nem történik számottevő változás ($p = 0{,}35$, U(alacsony mobilitás)=5). A problémát leíró döntési fa a 7.5. ábrán látható.



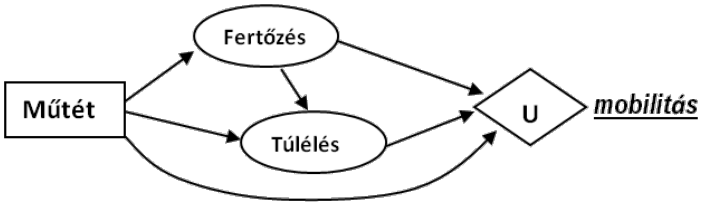
7.5. ábra. A térdműtét problémáját leíró döntési fa

A hasznosságcsomópontok értékei alapján kiszámítható a kapcsolódó valószínűségi csomópontok várható hasznossága (EU), ami felterjeszthető egészen a döntési csomópontig.

- EU(Szövődmény csomópont) = p(Szövődmény=igen) $\cdot$ U(halál) + + p(Szövődmény=nem) $\cdot$ U(rossz mobilitás) = $0{,}07 \cdot 0 + 0{,}93 \cdot 2 = 1{,}86$
- EU(Gyógyulás csomópont) = p(Gyógyulás=igen) $\cdot$ U(normális mobilitás) + + p(Gyógyulás=nem) $\cdot$ U(alacsony mobilitás) = $0{,}35 \cdot 5 + 0{,}65 \cdot 10 = 8{,}25$
- EU(Fertőzés csomópont) = p(Fertőzés=igen) $\cdot$ EU(Szövődmény csomópont) + + p(Fertőzés=nem) $\cdot$ EU(Gyógyulás csomópont)= $0{,}045 \cdot 1{,}86 + 0{,}955 \cdot 8{,}25 =$ 7,962
- EU(Műtét) = p(Komplikáció =igen) $\cdot$ U(halál) + p(Komplikáció=nem) $\cdot$ EU(Fertőzés csomópont)= $0{,}01 \cdot 0 + 0{,}99 \cdot 7{,}962 = 7{,}883$

Ezen számítás szerint a műtét várható hasznossága 7,883, szemben a műtét mellőzése esetén fennálló alacsony mobilitással járó hasznossággal (5). Mindezen leegyszerűsített számítások alapján a műtét elvégzése tűnik racionális döntésnek.

A fenti döntési fának egy lehetséges megfelelője döntési háló formájában a következő. A döntési háló (7.6. ábra) egy döntési (*Műtét*), egy hasznossági (*Mobilitás*) és két valószínűségi csomópontot tartalmaz, melyek a *Fertőzés: {igen, nem}*, *Túlélés: {teljes gyógyulás, részleges gyógyulás, halál}* értékeket vehetik fel. A döntési háló alkalmazásának előnye, hogy a hálóstruktúra és a hasznosságfüggvény megfelelő kialakításával csökkenthető az elvégzendő számítások száma. Például minden olyan esetben, ahol a *Túlélés=halál* értéke szerepel, ott a várható hasznosság nulla az *U(halál)=0* miatt.



Állapot		P	U	EU
Ha Műtét = 0		1	5	5
Különben, ha *Túlélés != halál*				
Ha *Túlélés = részleges gyógyulás*	*Fertőzés = igen*	0,041	2	0,083
	Fertőzés = nem	0,331	5	1,655
Ha *Túlélés = teljes gyógyulás*	*Fertőzés = nem*	0,6145	10	6,145
				7,883

7.6. ábra. A térdműtét problémát leíró döntési háló és a hozzá tartozó hasznosságfüggvény

Az élfordítási algoritmust követve az első lépés a *Túlélés* csomópont eliminálása, mivel ennek nincs más gyermeke, csak a hasznosságcsomópont. Ekkor az egyes hasznosságok szorzódnak a csomópont értékeihez (*teljes gyógyulás, részleges gyógyulás, halál*) kapcsolódó feltételes valószínűségekkel.

- $U(Fertőzés{=}igen,\ Műtét){=}\sum_T p(Túlélés|Fertőzés = igen,\ Műtét){\cdot}U(Túlélés,\ Fertőzés,\ Műtét)$
- $U(Fertőzés{=}igen,\ Műtét) = p(T{=}teljes\ gyógyulás|Fertőzés{=}igen,\ Műtét){\cdot}\ U(T{=}teljes\ gyógyulás,\ Fertőzés{=}igen,\ Műtét) + p(T{=}részleges\ gyógyulás|\ Fertőzés{=}igen,\ Műtét){\cdot}\ U(T{=}részleges\ gyógyulás,\ Fertőzés{=}igen,\ Műtét) + {+}p(T{=}halál|Fertőzés{=}igen,\ Műtét){\cdot}\ U(T{=}halál,\ Fertőzés{=}igen,\ Műtét)$
- $U(Fertőzés{=}igen,\ Műtét) = 0 \cdot 0 + (0{,}99 \cdot 0{,}93) \cdot 2 + (0{,}99 \cdot 0{,}07) \cdot 0 = 1{,}8414$
- $U(Fertőzés{=}nem,\ Műtét) = (0{,}99 \cdot 0{,}65) \cdot 10 + (0{,}99 \cdot 0{,}35) \cdot 5 + 0{,}01 \cdot 0 = 8{,}1675$

Ezt követi a *Fertőzés* csomópont eliminálása, majd az előző lépéshez hasonlóan kerül sor az új hasznosságok kiszámítására, ezúttal a *Fertőzés:{igen, nem}* értékekhez tartozó feltételes valószínűségekkel.

- $U(Műtét){=}\sum_F p(Fertőzés\ |Műtét){\cdot}U(Fertőzés,\ Műtét)$
- $U(Fertőzés{=}igen,\ Műtét) = p(Fertőzés{=}igen|\ Műtét){\cdot}U(Fertőzés{=}igen,\ Műtét) + p(Fertőzés{=}nem|\ Műtét){\cdot}U(Fertőzés{=}nem,\ Műtét)$

- $\mathrm{U}(\textit{Műtét}) = (0{,}045 \cdot 1{,}8414) + (0{,}955 \cdot 8{,}1675) = 7{,}883$

Ezen a ponton már rendelkezésre áll a végeredmény, a *Műtét=igen* várható hasznossága 7,883, míg a *Műtét=nem* esetében 5 a várható hasznosság. A végső lépésben a *Műtét* döntési csomópont eliminálására kerül sor, ekkor a maximális várható hasznosságú alternatívát kell választani, ami jelen esetben a *Műtét=igen*.

7.8. Költség-haszon elemzés

A modern egészségügy egyik alapvető kérdése, hogy egy beavatkozás költsége milyen mértékben van összhangban a beavatkozás egészségi állapotot érintő pozitív hatásaival, más szóval a beavatkozás mennyire költséghatékony. Az alapvető cél természetesen az, hogy a lehető legtöbb ember egészségi állapotán lehessen javítani, azonban a korlátos erőforrások miatt adott esetben határokat kell szabni. Ez a kérdéskör számos etikai, jogi és szociális kérdést vet fel, a jelenlegi keretek között azonban csak mérnöki megközelítésben fogjuk a problémát vizsgálni. Az alapfeltevés szerint a rendelkezésre álló korlátos mennyiségű erőforrást a lehető „legjobban" kell elkölteni úgy, hogy egyfelől az érintettek közössége számára közel optimális legyen, másfelől az érintettek többsége jól járjon. Leegyszerűsítve az előbbi kritérium azt jelenti, hogy a lehető legtöbb érintett részesüljön megfelelő ellátásban, míg az utóbbi azt, hogy egy adott érintett a számára lehető legjobb ellátást kapja. A beavatkozások hatékonyság-költség viszonyának vizsgálata azt a célt szolgája, hogy a jóság mértékét megragadó mércéket definiáljunk, és ezen mennyiségek révén az egyes alternatívák összevethetőek legyenek egymással.
Megjegyezzük, hogy míg a magyar szaknyelv egyaránt alkalmazza a *költség-haszon elemzés (cost-benefit analysis)*, és a *költség-hatékonyság vizsgálat (cost-effectiveness analysis)* kifejezéseket az egészségügyi szakterületeken, addig az angol szaknyelv legtöbbször az utóbbit használja erre a célra.

7.8.1. A hatékonyság mérése

Ahhoz, hogy a beavatkozások hasznát (hatékonyságát) vizsgálhassuk, tisztában kell lennünk három alapvető fogalommal.

- *Efficacy – hatásosság.* Kvantitatív mennyiség, amely azt adja meg, hogy egy adott folyamat vagy tevékenység mennyire képes egy elvárt hatást elérni. Az összehasonlítási alap mindig előre definiált. Orvosi értelmezésben azt jelenti, hogy egy vizsgált kezelés milyen mértékű javulást vagy gyógyulást idéz elő ideális (laboratóriumi vagy kísérleti) körülmények között.

- *Effectiveness – hatékonyság.* Eredetileg nem kvantitatív mennyiség, amely azt fejezi ki, hogy egy adott tevékenység révén a kitűzött célt sikerül-e elérni vagy sem, azonban gyakran használják kvantitatív értelemben is. Alapesetben nem nyújt információt arról, hogy mi a viszonyítási alap. Orvosi környezetben azt jelenti, hogy egy adott kezelés mennyire jól működik a klinikai gyakorlatban.

- *Efficiency – hathatóság, hatásfok.* Kvantitatív mennyiség, amely általánosan azt jelenti, hogy egy adott cél elérése érdekében mennyire jól használták fel a rendelkezésre álló időt és erőforrásokat. Legfőképpen annak mérésére használják, hogy

egy folyamat vagy tevékenység során adott mennyiségű bemenet mellett milyen mennyiségű elvárt kimenet keletkezett. Orvosi értelemben például ilyen lehet az a szám, hogy egy adott kezelés hatására 100 betegből hány gyógyult meg. Jellemzően hányadosként vagy százalékos formájában adják meg.

A gyakorlatban gyakori e fogalmak felcserélése, egymással való helyettesítése. Lényeges különbség a hatásosság és a hatékonyság között van, mivel az előbbit annak mérésére használják, hogy egy új kezelés (gyógyszer) van-e olyan jó kísérleti körülmények között, mint egy már meglévő kezelés (gyógyszer), míg az utóbbi a klinikai gyakorlatban történő alkalmazás során a kezelés jósága mértékének kifejezésére szolgál. A hathatóság és a hatékonyság közötti különbséget talán leginkább úgy lehet megragadni, hogy ez előbbi azt fejezi ki, hogy jól végezzük-e az adott tevékenységet, míg az utóbbi azt fejezi ki, hogy a megfelelő tevékenységet végezzük-e.
Emellett számos más mérce létezik, melyek egy része az *összehasonlító hatékonyságvizsgálatok (comparative effectiveness research, CER)* kiértékelését segítik elő. Ez utóbbira példa az *ACCE* keretrendszer, amellyel diagnosztikai vizsgálatok (tesztek) hatékonysága elemezhető négy vizsgált mennyiség segítségével [33]:

- *Analitikus validitás (analytic validity).* Meghatározza, hogy a mérendő mennyiség mennyire jól mérhető, a mérés reprodukálható-e. (Egy genotípus például jól mérhetőnek tekinthető, míg a családi kórelőzmény rosszul mérhetőnek számít.)

- *Klinikai validitás (clinical validity).* Megadja, hogy a vizsgált elem mennyire asszociált az adott betegség kialakulásával. (Például egy adott antitest jelenléte milyen mértékben asszociált a vizsgált betegség kialakulásával.)

- *Klinikai haszon (clinical utility).* Ez a tényező átfogó kvalitatív értékelést tesz lehetővé a várható hasznokról, kockázatokról, illetve arról, hogy a diagnosztikai vizsgálat eredménye mennyire képes orientálni a terápiaválasztást. (Például egy adott biomarker jelenléte esetén az alkalmazható gyógyszeres terápia nem hatékony, ezért alkalmazása nem javasolt.)

- *Hozzáadott klinikai érték (added clinical value).* Ez az érték azt fejezi ki, hogy más alternatívákhoz képest mennyivel nyújt többet ez a vizsgálat a kockázatok felmérése terén.

7.8.2. A költség és a hatékonyság viszonya

Ezidáig a hatékonyságot kizárólag a gyógyítás aspektusából értelmeztük, azonban a gyakorlatban nem hagyható figyelmen kívül a beavatkozások költségvonzata sem. A hatékonyság költségek függvényében történő megadására szolgál a költség-hatékonyság és a hozzá kapcsolódó széles eszköztár.
A költség-haszon (hatékonyság) elemzés (*Cost-Effectiveness Analysis, CEA*) egyik alapvető mennyisége a *nettó pénzügyi haszon* (net monetary benefit, NMB), amely egy adott I_i beavatkozás esetén a következőképpen definiálható [30]:

$$NMB_{I_i}(\lambda) = \lambda \cdot e_i - c_i, \tag{7.18}$$

ahol e_i a hatékonyság (effectiveness), c_i a költség (cost), a λ pedig egy paraméter, amely a hatékonyságot képezi le oly módon, hogy az pénzügyileg összevethető legyen a költséggel. Összességében az NMB segítségével meghatározható egy küszöb, amelyet a döntéshozó adott egységnyi egészségjavulásért hajlandó áldozni. A korábban már ismertetett QUALY jól alkalmazható ebben a kontextusban.

Abban az esetben, ha a beavatkozások kimenetele determinisztikus, azaz a hasznon egyértelműen meghatározható, akkor a költség-haszon elemzés egyértelműen megadja λ adott értékei esetén az optimális beavatkozást [32]. Az elemzés folyamán mindazon alternatívák eliminálódnak, melyeket egy-egy másik alternatíva dominál, majd ezt követően azok, melyeket egy alternatíva páros (kiterjesztetten) dominál. A végső eredmény a fennmaradt alternatívák inkrementális költség-haszon rátájának (incremental cost-effectiveness ratio, $ICER$) számítása alapján dől el, amely a következőképpen számítható:

$$ICER(I_i, I_j) = \frac{(c_i - c_j)}{(e_i - e_j)}, \tag{7.19}$$

ahol I_i, I_j két összevetendő beavatkozás. Az $ICER_{i,j}$ tehát megadja, hogy mekkora költség jut egységnyi hatékonyság mennyiségre, egyben definiál egy küszöbértéket, amelyre igaz, hogy $NMB_j > NMB_i$ ha $\lambda < ICER_{i,j}$ és $NMB_i > NMB_j$ ha $\lambda > ICER_{i,j}$. Ha feltesszük, hogy $c_i > c_j$, akkor $ICER_{i,j}$ alatt az olcsóbb alternatíva, felette pedig a drágább alternatíva lesz a jobb választás.

Legtöbbször azonban a beavatkozások kimenetele többféle lehet, melyek adott valószínűséggel következnek be. Ennek megfelelően akár gyökeresen eltérő kezelési útvonalak definiálhatók egy-egy betegségre az adott paraméterek függvényében. Az ilyen esetben alkalmazandó nemdeterminisztikus költség-haszon elemzés kivitelezhető döntési fák alkalmazásával vagy döntési hálókkal. A már ismertetett előnyei miatt az utóbbit vizsgáljuk a továbbiakban.

Általánosságban elmondható, hogy a cél az adott paraméterek estén optimális beavatkozás kiválasztása. Ezt nagymértékben elősegíti, ha az egyes beavatkozásokhoz kiszámítható, hogy azok mely paramétertartományokban hatékonyak, illetve hatékonyabbak másoknál. Ennek formális meghatározására szolgál a költség-haszon partíció (cost-effectiveness partition, CEP) [32].

Definíció Adott n intervallumhoz tartozó költség-haszon partíciófelosztás CEP egy $(\Theta, \underline{C}, \underline{E}, \underline{I})$ négyessel adható meg, ahol $\Theta = \theta_1, \ldots, \theta_{n-1}$ paraméterküszöbök $n-1$ elemű halmaza, $\underline{C} = c_0, \ldots, c_{n-1}$ a költségek n elemű halmaza, $\underline{E} = e_0, \ldots, e_{n-1}$ a hatékonyság-mértékek n elemű halmaza, $\underline{I} = I_0, \ldots, I_{n-1}$ pedig a beavatkozások n elemű halmaza.

Tételezzük fel, hogy $\theta_0 = 0$, ekkor az első partíció a $((0, \theta_1), c_0, e_0, I_0)$ négyessel adható meg, ami azt jelenti, hogy ha $0 < \lambda < \theta_1$, azaz λ az első partícióhoz tartozó paraméterintervallumban vesz fel értéket, akkor az I_0 beavatkozás lesz a leghatékonyabb c_0 költséggel és e_0 hatékonyságmértékkel. Értelemszerűen a i-edik partíció: $(\theta_{i-1}, \theta i), c_{i-1}, e_{i-1}, I_{i-1}$.

Nemdeterminisztikus esetben, ha adott $X = x_1, \ldots, x_m$ valószínűségi változó és annak $P(x_i)$ eloszlása, akkor a költség-haszon partíciófelosztás: CEP a lehetséges CEP_{X_i} partíciófelosztások súlyozott átlagaként előállítható, ha a súlyozott átlag számítása elvégezhető a költségekre és a hatékonyságmértékekre is. Tehát például ha $K(\lambda)$ meg-

adja az adott λ értéhez tartozó optimális beavatkozás költségét, $K_{CEP}(\lambda)$ előállítható $K_{CEP_i}(\lambda)$ függvények súlyozott összegeként [32], azaz

$$\forall\lambda, \quad K_{CEP}(\lambda) = \sum_{i=1}^{m} P(x_i) \cdot K_{CEP_i}(\lambda). \tag{7.20}$$

Ez a tulajdonság azért lényeges, mert ennek köszönhetően a CEP-számítás integrálható a döntési háló kiértékelési metódusába.

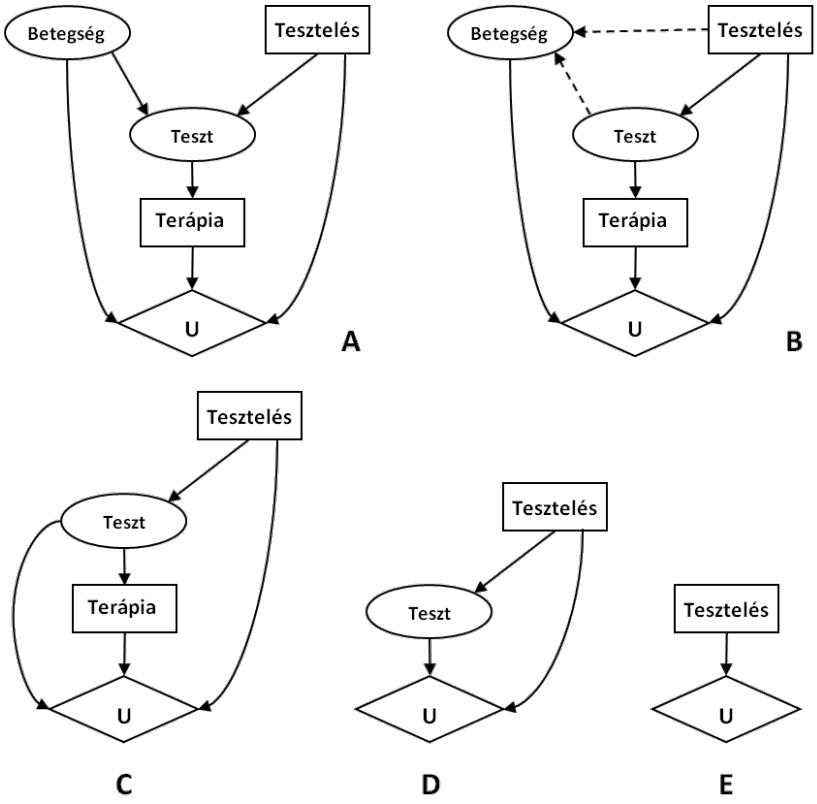
7.8.3. Költség-haszon elemzés mintapélda

Tekintsünk egy példát a döntési háló alapú költség-haszon elemzésre, amely költség-haszon partíció számítást is tartalmaz. Adott egy betegség, melynek a priori valószínűsége 0,08. Kifejlesztettek hozzá egy biomarkert, melynek jelenléte (pozitív teszteredmény) nagy valószínűséggel (0,94) jelzi a betegséget. A teszt specifikussága 0,89 (negatív teszteredmény a betegségtől való mentességre utal). A teszt elvégzése opcionális, költsége 200 euró. A páciens kétféle kezelésben részesülhet, melyek közül az egyik (T_1) relatíve olcsóbb (10 000 euró), közepes hatékonyságú (4), de nem káros az egészséges szervezetre (9,9), a másik (T_2) relatíve drágább (65 000 euró), hatékonyabb (8), viszont kismértékben káros az egészséges szervezetre (9). A harmadik eshetőség az, hogy a páciens nem részesül terápiában. Ekkor, ha a páciens beteg, nagymértékben leromlik az állapota (1). A kérdés az, hogy költséghatékony-e a teszt?
A problémát leíró döntési háló 5 csomópontból áll (7.7. ábra), melyek között egy összetett hasznosságcsomópont található, amely tartalmazza a költségeket és a hatékonyság mértékeket (7.1 táblázat). (Megjegyezzük, hogy lehetséges két külön hasznosságcsomópontként is modellezni a költségeket és hatékonyságot.)

7.1. táblázat. Költség-haszon elemzés mintapélda. A terápiák hatékonyság- és költségértékei

Kezelés	**Költség** EUR	**Hatékonyság** *beteg*	**Hatékonyság** *nem beteg*
Nincs terápia (NT)	0	1	10
1. terápia (T_1)	10 000	4	9,9
2. terápia (T_2)	65 000	8	9

A hasznosságcsomóponthoz kapcsolódik a *Betegség* valószínűségi csomópont, valamint a *Tesztelés* és a *Terápia* döntési csomópontok. A *Terápia* a terápiaválasztás révén befolyásolja a költségeket és a hatékonyságot, a *Tesztelés* pedig azt határozza meg, hogy sor kerüljön-e a teszt elvégzésére, s ezáltal hat a költségekre. Ha sor került a teszt elvégzésére, akkor annak kimenete befolyásolja a terápiaválasztást, amit a *Teszt* valószínűségi csomópont *Terápiához* való kapcsolódása jelez. A teszt eredménye nagymértékben függ attól, hogy a betegség valójában fennáll-e, így a *Betegség* is szülője a *Teszt* csomópontnak. A*Betegség* közvetlen kapcsolódása a hasznosságcsomóponthoz azt jelzi, hogy a betegség státusz alapvetően határozza meg a terápia hatékonyságát. A döntési háló kiindulási állapota a 7.7-A ábrán látható.



7.7. ábra. Döntési háló kiértékelésének lépései

A döntési háló kiértékelését a korábban már ismertetett élfodítás algoritmussal végezzük. Mivel zárványcsomópontok nincsenek, ezért először a hasznosságcsomóponthoz kapcsolódó döntési csomópontokat kell megvizsgálni, melyek akkor lennének eliminálhatók, ha minden más hasznosságcsomópontba mutató élű csomóponttal szülői kapcsolatban lennének. Ez sem a *Terápia*, sem a *Tesztelés* esetén nem teljesül, mivel a *Betegség*ből nem vezet él egyikbe sem. A következő lépésben a hasznosságcsomóponthoz kapcsolódó valószínűségi csomópontokat kell megvizsgálni. Jelen esetben ez a *Betegség*, amely akkor lenne eliminálható, ha máshová nem vezetne él belőle, csak a hasznosságcsomópontba. Ez sem teljesül, így ez a művelet sem alkalmazható. A negyedik és egyben utolsó lehetőség az élfordítás, mellyel valamelyik előbbi feltételnek megfelelő állapotot kell előállítani. Ennek megfelelően a *Betegség* $\rightarrow$ *Teszt* él megfordítását kell elvégeznünk. Az algoritmus szerint ekkor a *Teszt* szülei a *Betegség* szülei is lesznek, és vice versa. Tehát hozzáadunk egy *Tesztelés* $\rightarrow$ *Betegség* élt a hálóhoz, mivel az a *Teszt* szülője volt, a *Betegség*-nek azonban nem volt szűlője, így ezzel nincs teendő (7.7-B ábra). Ezt követően ki kell számolnunk a *Betegség* és a *Teszt* új feltételes valószínűségeit. A *Betegség*-nél lényegében a Bayes-szabályt alkalmazva állnak elő az új értékek, például:

- *P(beteg | Teszt="+", DT = 1)= P(Teszt="+" |beteg, DT = 1)* $\cdot$ *P(beteg) / P(Teszt="+" |beteg, DT = 1)*$\cdot$ *P(beteg) + P(Teszt="+" |nem beteg, DT = 1)*$\cdot$ *P(nem beteg)*, ahol $DT = 1$ a *Tesztelés = igen* döntést jelöli.

A *Teszt* csomópont esetében a korábbi feltételes valószínűségek összegzésével adódik

az új érték:

- *P(Teszt="+", DT) = P(Teszt="+" |beteg, DT) · P(beteg) + P(Teszt="+" |nem beteg, DT) · P(nem beteg).*

7.2. táblázat. A Betegség *és a* Teszt *csomópontok élfordítás előtti P(Teszt|Betegség, DT) feltételes valószínűségei. A DT a* Tesztelés *döntési csomópontot jelöli*

P(Teszt|Betegség, DT)	Beteg	Nem beteg
Teszt +	0,94	0,11
Teszt –	0,06	0,89
Nincs teszt	–	–

P(Beteg)	P(Nem beteg)
0,08	0,92

7.3. táblázat. A Betegség *és a* Teszt *csomópontok élfordítást követő P(Betegség|Teszt, DT) feltételes valószínűségei. A DT a* Tesztelés *döntési csomópontot jelöli*

P(B|Teszt,DT)	Teszt+	Teszt-	Nincs teszt
beteg	0,4263	0,0058	0,08
nem beteg	0,5737	0,9942	0,92

P(Teszt+|DT=1)	P(Teszt-|DT=1)
0,1764	0,8236

Az élfordítás előtti *P(Teszt|Betegség, DT)* és az azt követő *P(Betegség|Teszt, DT)* valószínűségeket rendre a 7.2 és a 7.3 táblázat összegzi. A tesztelés mellőzése esetén *DT = 0* speciális helyzet áll elő, ekkor a *Teszt* csomópont nem vesz fel értéket (*Nincs teszt*), vagyis a tőle függő *Betegség* csomópont az a priori valószínűségeknek megfelelő értékeket veszi fel. Az élfordítás következményeként a *Betegség* csomópont már eliminálható (7.7-C ábra), ekkor az hatékonyságértékek súlyozott átlagát kell képezni a megfelelő feltételes valószínűségekkel.

- *U(Teszt, Terápia, DT)* $= \sum_{betegseg} P(Betegség|Teszt, DT) \cdot U(Betegség, Terápia, DT)$, például:
- *U(Teszt="+", TR = nt, DT = 1) = P(beteg|Teszt="+", DT =1) · U(beteg, TR = nt, DT = 1) + P(nem beteg|Teszt="+", DT =1) · U(nem beteg, TR = nt, DT)*, ahol a *TR= nt* azt jelenti, hogy a páciens nem részesül terápiában. Konkrét értékekkel:
  - $U(Teszt="+",\ TR = nt,\ DT = 1) = 0{,}4263 \cdot 1 + 0{,}5737 \cdot 10 = 6{,}16,$
  - $U(Teszt="-",\ TR = nt,\ DT = 1) = 0{,}0058 \cdot 1 + 0{,}9942 \cdot 10 = 9{,}947,$
  - $U(Teszt=nincs,\ TR = nt,\ DT = 0) = 0{,}08 \cdot 1 + 0{,}92 \cdot 10 = 9{,}28.$

A hasznosságcsomópont tartalmát a számításokat követően az 7.4 táblázat foglalja össze. A *Betegség* csomópontból a hasznosságcsomópontba mutató élt pedig megörökli a *Teszt* csomópont.
Tovább folytatva a döntési háló kiértékelését a *Terápia* döntési csomópont eliminálása következik, mivel egyetlen gyermeke a hasznosságcsomópont, továbbá minden más csomópont, amely rendelkezik a hasznosságcsomópontba vezető éllel, egyúttal rendelkezik

7.4. táblázat. A hasznosságcsomópont hatékonyság- és költségértékei az egyes alternatívákra a Betegség *csomópont eliminálását követően*

	Tesztelés			Költség	Hatékonyság
1	D:Tesztelés	Van teszt	Teszt: +	200	6,16
2				10 200	7,38
3				65 200	8,57
4			Teszt: –	200	9,94
5				10 200	9,86
6				65 200	8,99
7		Nincs teszt	N/A	0	9,28
8				10 000	9,428
9				65 000	8,92

egy *Terápiá*-ba vezető éllel is (létezik egy irányított út, amelyen a *Terápia* elérhető). A *Teszt* csomópont közvetlen éllel bír a *Terápiá*-ba, a *Tesztelés* csomópontból pedig irányított úton elérhető a *Terápia*. Az utóbbi csomópont azonban továbbra sem kiértékelhető. A *Teszt* csomópont tehát megelőzi a terápiaválasztást (*Teszt*→*Terápia*), és egyben befolyásolja a hasznosságot is (*Teszt*→*U*). Ez azt jelenti, hogy a terápiaválasztás során mindenképpen figyelembe kell venni a *Teszt* lehetséges eredményeit (beleértve azt is, amikor nincs értéke). Tekintsük példaként azt esetet, amikor a tesztelésre sor került, és pozitív lett az eredménye (*DT = 1, Teszt="+"*). A 7.4 táblázat szerint ekkor három döntési alternatíva áll elő:

- *NT* azaz*Nincs terápia* ($c_1 = 200, e_1 = 6{,}16$),
- T_1 terápia ($c_2 = 10\,200, e_2 = 7{,}38$),
- T_2 terápia ($c_3 = 65\,200, e_3 = 8{,}57$).

Mivel egy döntési csomópont kiértékelésekor azt a döntést kell választani, amelyik a legnagyobb hasznosságú, itt meg kell vizsgálnunk, hogy a nettó pénzügyi haszon: $NMB_i = \lambda \cdot e_i - c_i$ szempontjából melyik alternatíva teljesít a legjobban. Ekkor láthatjuk, hogy egyértelműen jobb nincs, hanem adott tartományok között jobb az egyik, mint a másik. Ezeket a határokat a $\lambda \cdot 6{,}16 - 200 < \lambda \cdot 7{,}38 - 10\,200$, illetve a $\lambda \cdot 7{,}38 - 10\,200 < \lambda \cdot 8{,}57 - 65\,200$ egyenlőtlenségek kiértékelésével kapjuk meg. Ennek folyományaként két küszöb adódik: $\theta_1 = 8186{,}37$ az előbbi, $\theta_2 = 46\,261{,}70$ az utóbbi egyenlőtlenség révén. Ez tehát azt jelenti, hogy a legjobb döntés $0 < \lambda < \theta_1$ esetén a *Nincs terápia*, $\theta_1 < \lambda < \theta_2$ esetén a T_1 terápia, és $\theta_2 < \lambda$ esetén pedig T_2 terápia.
Abban az esetben, ha a teszt negatív volt (*DT = 1, Teszt="-"*), egyszerűbb helyzettel állunk szemben (lásd a 7.4 táblázat 4-6. sorait), ugyanis a *Nincs terápia* ($c_4 = 200, e_4 = 9{,}94$) dominálja a T_1 ($c_5 = 10\,200, e_5 = 9{,}86$) és a T_2 ($c_6 = 65\,200, e_6 = 8{,}99$) terápiákat, mivel minden λ-ra alacsonyabb költségű és nagyobb hatékonyságú. Ekkor tehát egyetlen tartomány van $(0, +\infty)$, ahol egyértelműen lehet dönteni.
Az utolsó eshetőség az, amikor nem kerül sor tesztelésre (*DT = 0, Teszt = nincs*). A három alternatíva közül (lásd a 7.4 táblázat 7-9. sorait) T_2-t kiterjesztetten dominálja T_1 és *Nincs terápia*, tehát T_2 elvethető. A másik két lehetőség közül egyik sem dominálja a másikat, így ismét meg kell vizsgálni a tartományokat. A $\lambda \cdot 9{,}28 - 0 < \lambda \cdot 9{,}428 - 10\,000$

elemzésével a $\theta_3 = 67\,567{,}60$ küszöb adódik, amely szerint $0 < \lambda < \theta_3$ esetén a *Nincs terápia* a legjobb döntés, $\theta_3 < \lambda$ esetén pedig a T_1 terápia ($c_8 = 10\,000, e_8 = 9{,}428$).

7.5. táblázat. Költség-haszon partíciók a Teszt *csomópont eliminálását követően*

$\theta_{i,j}$	θ_i	θ_j	**Költség**	**Hatékonyság**	**Terápia**	
$\theta_{0,1}$	0	8186,37	200	9,280	Teszt	NT
$\theta_{1,2}$	8186,37	46 261,70	1964	9,495	Teszt	-: NT
						+: T_1
$\theta_{2,inf}$	46 261,70	$+\infty$	11 666	9,705	Teszt	-: NT
						+: T_2
$\theta_{0,3}$	0	67 567,60	0	9,280	NT	
$\theta_{3,inf}$	67 567,60	$+\infty$	10 000	9,428	T_1	

A *Terápia* csomópont kiértékelését követően (7.7-D ábra) a *Teszt* csomópont kerül sorra, mivel egyetlen éle a hasznosságcsomópontba fut. Mivel a *Teszt* egy valószínűségi csomópont, ezért kiátlagolással (a hasznosság értékek *Teszt* szerinti súlyozott átlagolásával) eliminálunk. Ekkor a pozitív teszteredményekhez tartozó (teszten felüli) költségek és hatékonyság értékek a P(Teszt= "+") pozitív teszt valószínűségével szorzódnak (7.3táblázat), a negatív teszt esetében pedig P(Teszt= "-")-vel, majd ezután kerül sor tartományonkénti összegzésükre.

- $\theta_{0,1}$:(0, 8186,37)
  - Teszt="+": $c_1^+ = c_1 \cdot p(Teszt+|DT = 1) = 0 \cdot 0{,}1764 = 0, e_1^+ = e_1 \cdot p(Teszt+|DT = 1) = 6{,}16 \cdot 0{,}1764 = 1{,}0872$.
  - Teszt="-": $c_4^- = c_4 \cdot p(Teszt-|DT = 1) = 0 \cdot 0{,}8236 = 0, e_4^- = e_1 \cdot p(Teszt-|DT = 1) = 9{,}947 \cdot 0{,}8236 = 8{,}1928$.
  - $c_{\theta_{0,1}} = c_1^+ + c_4^- + c_{teszt} = 200$, $e_{\theta_{0,1}} = e_1^+ + e_4^- = 9{,}28$.
  - Ebben a tartományban sem pozitív, sem negatív teszt esetén nem kerül sor terápiára ($Terápia = NT$).
- $\theta_{1,2}$:(8186,37, 46 261,70)
  - Teszt="+": $c_2^+ = c_2 \cdot p(Teszt + |DT = 1) = 10\,000 \cdot 0{,}1764 = 1764, e_2^+ = e_2 \cdot p(Teszt + |DT = 1) = 7{,}38 \cdot 0{,}1764 = 1{,}3027$.
  - $c_{\theta_{1,2}} = c_2^+ + c_4^- + c_{teszt} = 1964$, $e_{\theta_{1,2}} = e_2^+ + e_4^- = 9{,}495$.
  - Ebben a tartományban pozitív teszt esetén T_1 terápiában részesül a beteg, negatív teszt esetén nem kerül sor terápiára.
- $\theta_{2,inf}$:(46 261,70, ∞)
  - Teszt="+": $c_3^+ = c_3 \cdot p(Teszt + |DT = 1) = 65\,000 \cdot 0{,}1764 = 11466, e_3^+ = e_3 \cdot p(Teszt + |DT = 1) = 8{,}57 \cdot 0{,}1764 = 1{,}5124$.
  - $c_{\theta_{2,inf}} = c_3^+ + c_4^- + c_{teszt} = 11\,666$, $e_{\theta_{2,inf}} = e_3^+ + e_4^- = 9{,}705$.
  - Ebben a tartományban pozitív teszt esetén T_2 terápiában részesül a beteg, negatív teszt esetén nem kerül sor terápiára.

Negatív teszt esetén a már említett okokból egyik tartományban sem kerül sor terápiára. Az adódó eredményeket a 7.5 táblázat foglalja össze, ahol látható, hogy azokat az eseteket nem érintette a súlyozás, ahol nem állt rendelkezésre teszt ($DT = 0$, azaz $\theta_{0,3}$ és $\theta_{3,inf}$).
A döntési háló kiértékelésének utolsó lépéseként a *Tesztelés* döntési csomópontot eliminálju k (7.7-D ábra). Ehhez össze kell vetnünk a teszt esetén és a teszt hiányában számított tartományokat, és minden esetben a lehető legjobb döntést kell hoznunk. Tehát a $\theta_{0,3}$ tartományt a $\theta_{0,1}, \theta_{1,2}, \theta_{2,inf}$ tartományokkal, illetve a $\theta_{3,inf}$ tartományt a $\theta_{2,inf}$-el. Ez a *Terápia* döntési csomópont kiértékeléséhez hasonlóan zajlik, azzal a különbséggel, hogy abban az esetben, ha nem létezik egyértelműen jobb választás, akkor $ICER$-számítással determinisztikus költség-haszon elemzést kell végezni. Továbbá fontos az a tény, hogy $ICER_{i,j}$ egy olyan küszöbérték, amelyre $c_i > c_j$ és $e_i > e_j$ esetén igaz, hogy alatta az olcsóbb alternatíva, felette pedig a drágább alternatíva lesz a jobb választás.

- ($\theta_{0,3}$ és $\theta_{0,1}$): a hatékonyságuk azonos, költség viszont $\theta_{0,3}$-ban $DT = 0$ esetén nincs, mivel nem kerül sor tesztelésre. Ebben a tartományban tehát $DT = 0$ a megfelelő választás, mivel ez dominálja a másikat.

- ($\theta_{0,3}$ és $\theta_{1,2}$): nincs egyértelműen jobb, ICER=9114,53, ami ketté osztja a tartományt. (8186,37 – 9114,53) között a $DT = 0$ a megfelelő döntés, (9114,53 – 46 261,70) között pedig a $\theta_{1,2}$ által diktált $DT = 1$, *Teszt -*: *nincs terápia*, *Teszt +*: T_1.

- ($\theta_{0,3}$ és $\theta_{2,inf}$): nincs egyértelműen jobb, ICER=27 436,5, ami a tartományon kívül esik így a $\theta_{2,inf}$ által diktált $DT = 1$ döntés a jobb, *Teszt -*: *nincs terápia*, *Teszt +*: T_2.

- ($\theta_{3,inf}$ és $\theta_{2,inf}$): nincs egyértelműen jobb, ICER=6010,101, ami a tartományon kívül esik így a $\theta_{2,inf}$ által diktált $DT = 1$ döntés a jobb, *Teszt -*: *nincs terápia*, *Teszt +*: T_2.

A számítások eredményeként előálló költség-hatékonysági tartományokat a 7.6 táblázat tartalmazza. Ezt figyelmesen megvizsgálva látható, hogy az első és az utolsó két tartomány összevonható, mivel pontosan ugyanaz a terápiaválasztás, költség és hatékonyság jellemzi őket. Ennek következtében végeredményben három CEP állt elő:

- (0 – 9114,53): nincs tesztelés ($DT = 0$), nincs terápia.

- (9114,53 – 46 261,70): van tesztelés ($DT = 1$), ha *Teszt +*: T_1 terápia, ha *Teszt -*: nincs terápia.

- (46 261,70 – ∞): van tesztelés ($DT = 1$), ha *Teszt +*: T_2 terápia, ha *Teszt -*: nincs terápia.

A feladat elején feltett kérdésre tehát azt válaszolhatjuk, hogy $\lambda > 9114{,}53$ EUR/QALY felett a teszt költséghatékony.
Mint ahogy az látható, még egy egyszerű döntési háló kiértékelése is meglehetősen komplex feladat, emiatt a gyakorlatban kizárólag szoftveres megoldások segítségével kerül sor a megtervezésükre, kiértékelésükre és finomhangolásukra.

7.6. táblázat. Költség-haszon partíciók

	θ_i	θ_{i+1}	**Tesztelés**	**Terápia**	**Költség**	**Hatékonyság**
1	0	8186,37	nem	NT	0	9,280
2	8186,37	9114,53	nem	NT	0	9,280
3	9114,53	46 261,70	igen	$T^- : NT, T^+ : T_1$	1964	9,495
4	46 261,70	67 567,60	igen	$T^- : NT, T^+ : T_2$	11 666	9,705
5	67 567,60	$+\infty$	igen	$T^- : NT, T^+ : T_2$	11 666	9,705

Irodalomjegyzék

[1] Tversky, A. and Kahneman, D. *Belief in the law of small numbers* in Kahneman, D. and Slovic, P. and Tversky, A., editors: *Judgment under Uncertainty: Heuristics and Biases*, pages 23-31. Cambridge University Press, New York, NY, 1982.

[2] A. Tversky and D. Kahneman. *Causal schemas in judgements under uncertainty.* In D. Kahneman, P. Slovic, and A. Tversky, editors, *Judgment under Uncertainty: Heuristics and Biases*, pages 117-128. Cambridge University Press, New York, NY, 1982.

[3] D Kahneman and A. Tversky. *Subjective probability: A judgement of representativeness, 1982.*

[4] D Kahneman and A. Tversky. *Intuitive prediction: Biases and corrective procedures.* In D. Kahneman, P. Slovic, and A. Tversky, editors, Judgment under Uncertainty: *Heuristics and Biases*, pages 414-421. Cambridge University Press, New York, NY, 1982.

[5] S. Lichtenstein, B. Fischhoff, and L. D. Phillips. *Calibration of probabilities.* In D. Kahneman, P. Slovic, and A. Tversky, editors, *Judgment under Uncertainty: Heuristics and Biases*, pages 306-334. Cambridge University Press, New York, NY, 1982.

[6] S. Lichtenstein, B. Fischhoff, and L. D. Phillips. *Calibration of probabilities.* In D. Kahneman, P. Slovic, and A. Tversky, editors, *Judgment under Uncertainty: Heuristics and Biases*, pages 335-354. Cambridge University Press, New York, NY, 1982.

[7] T. Englander. *Viaskodás a bizonytalannal: A valószínűségi ítéletalkotás egyes pszichológiai problémái.* Akadémiai kiadó, Budapest, 1999.

[8] M. J. Druzdzel and A. Onisko. *The impact of overconfidence bias on practical accuracy of bayesian network models: An empirical study.* In *Working Notes of the 2008 Bayesian Modelling Applications Workshop, Special Theme: How Biased Are Our Numbers? Part of the Annual Conference on Uncertainty in Artificial Intelligence (UAI-2008), 2008.*

[9] A. Onisko and M. J. Druzdzel. *Impact of quality of bayesian network parameters on accuracy of medical diagnostic systems. In AIME'11 Workshop on Probabilistic Problem Solving in Biomedicine (ProBioMed-11), 2011.*

[10] Leslie A McArthur. *The how and what of why: Some determinants and consequences of causal attribution. Journal of Personality and Social Psychology, 22(2):171-193, 1972.*

[11] Harold H Kelley. *The processes of causal attribution. American Psychologist, 28(2):107-128, 1973.*

[12] P. Tang. *Key capabilities of an electronic health record system.* In A. Philip, J. M. Corrigan, J. Wolcott, and S. M. Erickson, editors, *Patient Safety: Achieving a New Standard for Care. Institute of Medicine Committee on Data Standards for Patient Safety. Board on Health Care Services., pages 430-467.* National Academies Press, Washington D.C., 2003.

[13] W. J. Clancey and E. H. Shortliffe. *Medical artificial intelligence programs.* In W. J. Clancey and E. H. Shortliffe, editors, *Readings in Medical Artificial Intelligence: The First Decade.* AAAI, 1984.

[14] M.G. Kahn, S. A. Steib, V. J. Fraser, and W. C. Dunagan. *An expert system for culture-based infection control surveillance.* In *Proceedings of the Annual Symposium on Computer Applications in Medical Care.*, pages 171-175, 1993.

[15] G. Edwards, P. Compton, R. Malor, A. Srinivasan, and L. Lazarus. *Peirs: a pathologist maintained expert system for the interpretation of chemical pathology reports.* Pathology, 25(1):27-34, 1993.

[16] G. O. Barnett, J. J. Cimino, J. A. Hupp, and E. P. Hoffer. *Dxplain. an evolving diagnostic decision-support system.* JAMA, 258(1):67-74, 1987.

[17] R.M. Gardner, T. A. Pryor, and H. R. Warner. *The help hospital information system: update 1998.* Int. Journal of Medical Informatics, 54(3):169-182, 1999.

[18] I. Bratko, I. Mozetic, and N. Lavrac. *KARDIO: A Study in Deep and Qualitative Knowledge for Expert Systems.* MIT Press, Cambridge, MA, 1989.

[19] J. Wyatt and D. Spiegelhalter. *Field trials of medical decision-aids: potential problems and solutions.* In Proceedings of the 15th Symposium on Computer Applications in Medical Care., pages 3-7. McGraw Hill Inc., 1991.

[20] L. Perreault and J. Metzger. *A pragmatic framework for understanding clinical decision support.* Journal of Healthcare Information Management., 13(2):5-21, 1999.

[21] J. Neumann and O. Morgenstern. *Theory of Games and Economic Behavior.* Princeton University Press, Princeton, NJ, 1944.

[22] S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach. 2nd edition.* Prentice Hall, New Jersey, 2002.

[23] S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach. 3rd edition.* Prentice Hall, New Jersey, 2009.

[24] P. Kind, R. Brooks, and R. Rabin. *EQ-5D concepts and methods.* Springer, http://www.euroqol.org/,2005.

[25] J. S. Pliskin, D. S. Shepard, and M. C. Weinstein. *Utility functions for life years and health status.* Operations Research., 28(1):206-224, 1980.

[26] R.D. Shachter. *Evaluating influence diagrams. Operations Research.*, 34(6):871-882, 1986.

[27] NICE. Measuring effectiveness and cost effectiveness: the QALY. National Institute of Health and Clinical Excellence, UK., http://www.nice.org.uk/newsroom/features/measuringeffectivenessandcosteffectivenesstheqaly.jsp,2010.

[28] D. Spiegelhalter and M. Pearson. *Understanding uncertainty: Small but lethal. PLUS.* http://plus.maths.org/content/os/issue55/features/risk/index,55,2010.

[29] R. A. Howard. On making life and death decisions. In J. Richard, C. Schwing, and W. A. Albers, editors, *Societal Risk Assessment: How Safe Is Safe Enough?* General Motors Research Laboratories. Plenum Press, New York, 1980.

[30] A. A. Stinnett and J. Mullahy. *Net health benefit: A new framework for the analysis of uncertainity in cost-effectiveness analysis.* Medical Decision Making, 18(2S):68-80, 1998.

[31] S. M. Olmsted. On *Representing and Solving Decision Problems. PhD thesis. Dept. Engeneering-Economic Systems*, Stanford University., CA, 1983.

[32] M. Arias and F. J. Diez. *Cost-effectiveness analysis with influence diagrams.* In AIME11 Workshop on Probabilistic Problem Solving in Biomedicine. Bled,Slovenia., pages 121-135. ProBioMed11, 2011.

[33] K.A. Goddard, W.A. Knaus, E. Whitlock, G.H. Lyman, H.S. Feigelson, S.D. Schully, S. Ramsey, S. Tunis, A.N. Freedman, M.J. Khoury, and D.L. Veenstra. *Building the evidence base for decision making in cancer genomic medicine using comparative effectiveness research.* Genetics in Medicine, 14(7):633-642, 2012.

[34] M. Forstner. *Benefit-–risk management in the age of personalized healthcare. Personalized Medicine*, 9(5):507-514, 2012.

[35] R.M. Wenham, D.M. Sullivan, M. Hulse, P.B. Jacobsen, and W.S. Dalton. *The creation of an integrated health-information platform: building the framework to support personalized medicine. Personalized Medicine*, 9(6):621-632, 2012.

[36] D.K. Owens, R.D. Shachter, and R.F.Jr. Nease. *Representation and analysis of medical decision problems with influence diagrams.* Medical Decision Making, 17(3):241-262, 1997.

[37] OpenClinical.org. *Openclinical: Knowledge management for medical care*, 2012.