Valószínűségszámítás jegyzet 4. rész

1. Nagy számok törvényei: Csebisev-, Bernoulli- és Kolmogorov-féle alak. Centrális határeloszlás-tétel, de Moivre–Laplace-tétel.

Ebben a fejezetben két nagyon fontos témakört érintünk. Az első az átlag és a várható érték kapcsolatát formalizálja, míg a második az átlag eloszlását vizsgálja.

1.1. Nagy számok törvényei

A most következő három tétel egyre erősebb összefüggést ad, azaz visszafele tekintve egymás következményei. Mi időrendi sorrendben nézzük őket, mert így alapozták meg egymásnak az utat. Innentől kezdve az NSzT rövidítést használjuk a nagy számok törvénye szókapcsolat helyett.

1. Tétel. *(Bernoulli-féle NSzT) Tegyük fel, hogy egy kísérletet egymástól teljesen függetlenül sokszor elvégzünk, és X_i az i-edik kísérleten az A esemény indikátora. Ekkor bármely $\varepsilon > 0$ valós számra igaz, hogy*

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - P(A)\right| \geq \varepsilon\right) = 0.$$

Azaz röviden megfogalmazva az indikátorok átlaga tart az indikátorok várható értékéhez (hiszen $P(A)$ a várható értéke az indikátor valószínűségi változónak). Azonban a valószínűségszámításban több különböző konvergencia fogalom létezik, amik sajnos nem ekvivalensek. A fenti tételben szereplő tulajdonság az úgynevezett **sztochasztikus konvergenciája** az átlagnak a várható értékhez.

Az alábbi tétel megmutatja, hogy az 1. Tétel kiterjeszthető nem indikátor valószínűségi változókra, és nem szükséges a teljesen függetlenség.

2. Tétel. *(Csebisev-féle NSzT) Tegyük fel, hogy az $X_1, X_2, ..., X_n, ...$ azonos eloszlású, páronként független valószínűségi változóknak létezik EX várható értékük és σX szórásuk. Ekkor bármely $\varepsilon > 0$ valós számra igaz, hogy*

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - EX\right| \geq \varepsilon\right) = 0.$$

Bizonyítás: A határértékben szereplő valószínűséget a Csebisev-egyenlőtlenséggel fogjuk becsülni. Vegyük észre, hogy $E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right) = EX$*. A képlethez már csak az átlag szórásnégyzetére van szükségünk:*

$$\sigma^2\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right) = \frac{1}{n^2}\sigma^2\left(\sum_{i=1}^{n}X_i\right),$$

ami a páronkénti függetlenség következtében

$$\frac{1}{n^2}\sum_{i=1}^{n}\left(\sigma^2 X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\left(\sigma^2 X\right) = \frac{\sigma^2 X}{n}.$$

A Csebisev-egyenlőtlenséget felírva:

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n}X_i - EX\right| \geq \varepsilon\right) \leq \frac{\frac{\sigma^2 X}{n}}{\varepsilon^2}.$$

Erről a felső korlátról könnyen látható, hogy 0 *hoz tart, ahogy* n *tart végtelenhez, így a kérdéses valószínűségnek is* 0*-hoz kell tartania.*

Tehát az átlag sztochasztikusan konvergál a várható értékhez. Azonban létezik a sztochasztikusnál erősebb konvergencia fogalom is, erről szól az alábbi tétel:

3. Tétel. *(Kolmogorov-féle NSzT) Tegyük fel, hogy az* $X_1, X_2, ..., X_n, ...$ *teljesen független valószínűségi változók, amik várható értéke megegyezik (*EX*), és szórásaikra igaz, hogy* $\sum_{i=1}^{\infty}\frac{1}{i^2}\sigma^2 X_i < \infty$*. Ekkor igaz, hogy*

$$P\left(\lim_{n\to\infty}\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right) = EX\right) = 1.$$

A tételben szereplő konvergenciát **1 valószínűségű konvergencia**ként említik. Ahogy a tétel előtt már utaltunk rá, ez egy erősebb konvergencia fogalom, mint a sztochasztikus.

Nézzünk egy példát, amikor egy valószínűségi változó sorozat sztochasztikusan konvergens, de 1 valószínűséggel nem: Legyen $X \in U(0,1)$ *valószínűségi változóból generálva az* X_i *indikátor valószínűségi változó sorozat a következő módon:* X_1 *annak az indikátora, hogy* $X \in [0, \frac{1}{2}]$*;* X_2 *annak, hogy* $X \in (\frac{1}{2}, 1]$*;* X_3 *annak, hogy* $X \in [0, \frac{1}{3}]$*;* X_4 *annak, hogy* $X \in (\frac{1}{3}, \frac{2}{3}]$*;* X_5 *annak, hogy* $X \in (\frac{2}{3}, 1]$*;* X_6 *annak, hogy* $X \in [0, \frac{1}{4}]$*... Azaz* k *egyenlő hosszú részre osztjuk az intervallumot, majd mindegyik résznek sorra vesszük az indikátorát, ezután folytatjuk* 1*-gyel nagyobb* k *értékre. Az így kapott valószínűségi változó sorozat sztochaszikusan konvergál a konstans 0-hoz, hiszen ahogy* n *tart végtelenhez, egyre nagyobb szünet jön két* 1 *értékű között. Azonban bármely* X *értékre és* N *számra létezik* $i > N$*, hogy* $X_i = 1$*, azaz a sorozat nem konvergál* 0*-hoz 1 valószínűséggel.*

1.2. Centrális határeloszlás-tételek

Amikor bevezettük a normális eloszlást, már szóba került, hogy származtathatjuk a binomiális eloszlásból, ha kivonjuk a várható értéket, és leosztunk n-el. Valójában a standardizáltjára van szükségünk, ahogy az alábbi tétel is mutatja:

4. Tétel. *(de Moivre–Laplace-tétel) Tegyük fel, hogy egy kísérletet egymástól teljesen függetlenül sokszor elvégzünk, és X_i az i-edik kísérleten az A esemény indikátora. Ekkor bármely x valós számra igaz, hogy*

$$\lim_{n\to\infty} P\left(\frac{\sum_{i=1}^{n} X_i - n\cdot P(A)}{\sqrt{nP(A)(1-P(A))}} < x\right) = \Phi(x).$$

Arról már korábban is beszéltünk, hogy karakterisztikusok összegeként binomiális eloszlást kapunk. Ez a tétel azonban azt mondja, hogy ha ezt az összeget standardizáljuk, akkor a határérték standard normális eloszlású lesz. Vegyük észre, hogy valóban standardizáltról van szó, hiszen $E(\sum_{i=1}^{n} X_i) = n\cdot P(A)$ és $\sigma^2(\sum_{i=1}^{n} X_i) = nP(A)(1-P(A))$. A tételben szereplő konvergencia is különleges szereppel bír a valószínűségszámításban: azt mondjuk, hogy az indikátor valószínűségi változók összegének standardizáltja **eloszlásban konvergál** a standard normálishoz. Ezáltal persze kaphatunk egy közelítést a Φ függvényre:

5. Következmény. *Ha $0 < p < 1$ és $q = 1-p$, akkor*

$$\lim_{n\to\infty} \sum_{i=0}^{i<x\sqrt{npq}+np} \binom{n}{i} p^i q^{n-i} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt.$$

Bizonyítás: Legyen $p = P(A)$ és $Y \in Bin(n,p)$. A de Moivre–Laplace-tételben szereplő valószínűséget kell csak átírnunk:

$$P\left(\frac{\sum_{i=1}^{n} X_i - n\cdot P(A)}{\sqrt{nP(A)(1-P(A))}} < x\right) = P\left(\frac{Y - n\cdot p}{\sqrt{n\cdot p\cdot q}} < x\right) =$$

$$= P\left(Y < x\sqrt{n\cdot p\cdot q} + n\cdot p\right) = \sum_{i=0}^{i<x\sqrt{npq}+np} \binom{n}{i} p^i q^{n-i}.$$

A de Moivre–Laplace-tételben szereplő konvergencia sebességéről itt mi nem beszélünk, de tipikusan százas nagyságrendű kísérletszám esetén már jó közelítést kaphatunk a binomiális eloszlást követő valószínűségi változóra, ha a tétel szerinti Φ értékekből következtetünk.

Mennyi annak a valószínűsége, hogy 100 véletlenszerűen választott emberből kevesebb, mint 55 nő?
A kérdéses X szám értékre $X \in Bin(100, \frac{1}{2})$. Így a standardizáltja a de Moivre–Laplace-tétel értelmében standard normálishoz közeli. Így

$$P(X < 55) = P\left(\frac{X - 50}{\sqrt{25}} < \frac{5}{\sqrt{25}}\right) \approx \Phi(1) \approx 0{,}8413.$$

Ami ennél sokkal meglepőbb lehet az az, hogy nem csak az indikátorok összegének standardizáltja tart a standard normális eloszláshoz:

6. Tétel. *(Centrális határeloszlás-tétel) Tegyük fel, hogy az $X_1, X_2, ..., X_n, ...$ azonos eloszlású, teljesen független valószínűségi változóknak létezik EX várható értékük és σ szórásuk. Ekkor:*

$$\lim_{n\to\infty} P\left(\frac{1}{\sigma\sqrt{n}}\left(\sum_{i=1}^{n} X_i - n \cdot EX\right) < x\right) = \Phi(x).$$

Tehát a centrális határeloszlás-tétel (röviden CHT) értelmében bármilyen eloszlásból teljesen függetlenül választott sorozat összegének standardizáltja eloszlásban tart a standard normálishoz. Ahogyan a de Moivre–Laplace-tételnél, itt sem beszélünk a tételben szereplő konvergencia sebességéről, de megemlítjük, hogy százas nagyságrendű kísérletszám esetén már jó közelítést kaphatunk az összeg standardizáltjára, ha a tétel szerinti Φ értékekből következtetünk. (A konvergencia sebességére a Berry–Essen-tétel ad korlátot.)

Hogy érzékeltessük a CHT jelentőségét, különböző eloszlású valószínűségi változók összegére nézünk alább példát. A konvolúciónál tanult nevezetes összegeknél tanultak alapján pontosan tudjuk, hogy azonos eloszlású, teljesen független binomiálisok összege, Poisson-eloszlásúak összege illetve normális eloszlásúak összege miként viselkedik. Ezekre az eloszlásokra ezért nem nézünk külön példát.

1. *Egy célbalövés sorozat akkor ér véget, ha 100 alkalommal eltalálta a versenyző a célt. Mennyi annak a valószínűsége, hogy egy 75%-os találati aránnyal lövő játékosnak kevesebb, mint 150 lövésre lesz szüksége?*
   Az i-edik találathoz használt töltények száma $X_i \in G(0{,}75)$. A CHT értelmében (mivel $EX_i = \frac{4}{3}$ és $\sigma X_i = \frac{2}{3}$):

$$P\left(\sum_{i=1}^{100} X_i < 150\right) = P\left(\frac{3}{20}\left(\sum_{i=1}^{100} X_i - \frac{400}{3}\right) < \frac{3}{20}\left(150 - \frac{400}{3}\right)\right) \approx$$

$$\approx \Phi\left(\frac{3}{20}\left(150 - \frac{400}{3}\right)\right) = \Phi(2{,}5) \approx 0{,}9938.$$

2. *100-szor egyenletesen választunk egy-egy számot a $(0,10)$ intervallumon. Mennyi annak a valószínűsége, hogy a választott számok összege kevesebb, mint 550?*
   Az i-edik szám eloszlása $X_i \in U(0,10)$. A CHT értelmében (mivel $EX_i = 5$ és $\sigma X_i = \frac{5}{\sqrt{3}}$):

$$P\left(\sum_{i=1}^{100} X_i < 550\right) = P\left(\frac{\sqrt{3}}{50}\left(\sum_{i=1}^{100} X_i - 500\right) < \frac{\sqrt{3}}{50}(550-500)\right) \approx$$

$$\approx \Phi\left(\frac{\sqrt{3}}{50}(550-500)\right) = \Phi(\sqrt{3}) \approx 0{,}9582.$$

3. *Egy fogaskerék első meghibásodásának ideje örökifjú tulajdonságú. Várhatóan 1 hét után kell először cserélni. Mennyi annak a valószínűsége, hogy egy 100-as csomag fogaskereket kevesebb, mint 110 hét alatt fel kell használnunk?*
   Az i-edik fogaskerék használatának időtartama $X_i \in Exp(1)$. A CHT értelmében ($EX_i = 1$ és $\sigma X_i = 1$):

$$P\left(\sum_{i=1}^{100} X_i < 110\right) = P\left(\frac{1}{10}\left(\sum_{i=1}^{100} X_i - 100\right) < \frac{1}{10}(110-100)\right) \approx$$

$$\approx \Phi\left(\frac{1}{10}(110-100)\right) = \Phi(1) \approx 0{,}8413.$$

4. *Egy bárban a napi sörfogyasztás a csapról hektoliterekben mérve X, amire ismerjük, hogy $f_X(t) = 2 - 3t^2$, ahol $0 < t < 1$. Mennyi annak a valószínűsége, hogy elég lesz a sör, ha tudjuk, hogy 26 hektoliter sör van a bárban, de 100 napig nem érkezik új szállítmány?*
   Legyen az i-edik napi fogyasztás mértéke X_i. Ekkor

$$EX_i = \int_0^1 2t - 3t^3 dt = \left[t^2 - \frac{3}{4}t^4\right]_0^1 = \frac{1}{4},$$

$$E(X_i^2) = \int_0^1 2t^2 - 3t^4 dt = \left[\frac{2}{3}t^3 - \frac{3}{5}t^5\right]_0^1 = \frac{1}{15},$$

$$\sigma^2 X_i = E(X_i^2) - E(X_i)^2 = \frac{1}{15} - \frac{1}{16} = \frac{1}{240}.$$

   Így a CHT értelmében:

$$P\left(\sum_{i=1}^{100} X_i < 26\right) = P\left(\frac{\sqrt{240}}{10}\left(\sum_{i=1}^{100} X_i - 25\right) < \frac{\sqrt{240}}{10}(26-25)\right) \approx$$

$$\approx \Phi\left(\frac{\sqrt{240}}{10}(26-25)\right) = \Phi(1{,}5492) \approx 0{,}9394.$$

2. Feltételes eloszlás, feltételes várható érték, lineáris regresszió. A regresszió tulajdonságai. Példák diszkrét és folytonos esetben.

Egy eseményre feltételezetten vizsgálni egy valószínűségi változót magától értetődő, egyszerűen minden valószínűséget a feltételezett eseményre megszorított eseménytéren értelmezett valószínűségi mezőn vizsgálunk. Mi a helyzet, ha nem eseményre, hanem valószínűségi változóra akarunk feltételezni?

7. Definíció. *Az X valószínűségi változó Y valószínűségi változóra való feltételes eloszlásfüggvénye az y paraméter függvényében*

$$F_{X|Y}(x|y) = P(X < x|Y = y).$$

Abban az esetben, ha Y eloszlása folytonos, az $Y = y$ esemény 0 valószínűségű, így az eredeti definíció szerint nem feltételezhetünk rá. Emlékezzünk, hogy a folytonos teljes valószínűség tételnél pontosan ezért már kiterjesztettük a feltételes valószínűség definícióját. A fenti definícióba behelyettesítve a feltételes velószínűség általános képletét a következőt kapjuk:

8. Következmény. *Az X valószínűségi változó Y valószínűségi változóra való feltételes eloszlásfüggvénye*

$$F_{X|Y}(x|y) = \lim_{\varepsilon \to 0} P(X < x|y \leq Y < y + \varepsilon).$$

A feltételes eloszlás persze minden paraméter esetén egy-egy eloszláshoz vezet. Ezen eloszlások vizsgálata során attól függően, hogy X vagy Y diszkrét eloszlásúak-e, más metodikát alkalmazunk (szumma vagy integrál; valószínűségek vagy sűrűségfüggvény).

Nézzünk meg egy diszkrét és egy folytonos példát is. Diszkrét esetben a valószínűségek megadásával írjuk le az eloszlást, míg folytonos esetben az eloszlásfüggvénnyel.

1. *Dobunk egy szabályos kockával, majd az eredménynek megfelelő számú szabályos érmével. Legyen az Y valószínűségi változónk a kockadobás értéke, X pedig a fejek száma az érmedobások között. Mi lesz az $X|Y$ eloszlása?*

$$P(X = x|Y = y) = \left(\frac{1}{2}\right)^y \binom{y}{x} \text{ , ahol } x = 0, 1, ..., y,$$

   hiszen ekkor $X \in Bin(y, \frac{1}{2})$. Figyeljünk oda, hogy X eloszlása alapvetően nem binomiális, csak rögzített y-ra feltételezve az.

2. *Eltörünk egy 1 méter hosszú pálcát egy, a középső egyharmad méteren egyenletesen választott pontban. A rövidebbik pálcán megismételjük a műveletet (azaz annak a középső harmadán is egyenletesen törünk). Legyen az Y valószínűségi változónk az első töréspont helye, azaz $Y \in U(\frac{1}{3}, \frac{2}{3})$. Legyen továbbá X a legrövidebb kapott pálca hossza. Mi lesz az $X|Y$ eloszlása?*
   Ha az $Y = y \leq \frac{1}{2}$, akkor az y hosszú darab lesz tovább törve, így:

$$F_{X|Y}(x|y) = P(X < x|Y = y) = \frac{x - \frac{y}{3}}{\frac{y}{6}}\ , \text{ ahol } \frac{y}{3} \leq x \leq \frac{y}{2},$$

   hiszen $\frac{y}{3}$ és $\frac{y}{2}$ között egyenletesen helyezkedhet el X.
   Ha az $Y = y > \frac{1}{2}$, akkor az $1-y$ hosszú darabot törjük tovább, így:

$$F_{X|Y}(x|y) = P(X < x|Y = y) = \frac{x - \frac{1-y}{3}}{\frac{1-y}{6}}\ , \text{ ahol } \frac{1-y}{3} \leq x \leq \frac{1-y}{2},$$

   hiszen $\frac{1-y}{3}$ és $\frac{1-y}{2}$ között egyenletesen helyezkedhet el X.

9. Tétel. *Ha X és Y együttes eloszlása folytonos, akkor*

$$F_{X|Y}(x|y) = \frac{\int_{-\infty}^{x} f_{X,Y}(t,y)dt}{f_Y(y)} = \frac{\frac{d(F_{X,Y}(x,y))}{dy}}{f_Y(y)}$$

Ezzel tehát az együttes eloszlásfüggvényből számítható a feltételes eloszlásfüggvény. Ebből pedig már rögtön adódik a feltételes sűrűségfüggvény is:

10. Tétel. *Ha X és Y együttes eloszlása folytonos, akkor*

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{\frac{d^2(F_{X,Y}(x,y))}{dx \cdot dy}}{f_Y(y)}$$

Legyen az X, Y valószínűségi változók együttes sűrűségfüggvénye $f_{X,Y}(x,y) = 15xy^2$, ahol $0 < x < 1$ és $0 < y < x$. Határozzuk meg az $X|Y$ feltételes eloszlás eloszlásfüggvényét és sűrűségfüggvényét!
A 9. tételben látott képlet szerint:

$$F_{X|Y}(x|y) = \frac{\int_{-\infty}^{x} f_{X,Y}(t,y)dt}{f_Y(y)},$$

amibe behelyettesítve az együttes sűrűségfüggvényt (a határokra odafigyelve):

$$F_{X|Y}(x|y) = \frac{\int_{y}^{x} 15ty^2dt}{f_Y(y)} = \frac{\left[\frac{15}{2}t^2y^2\right]_y^x}{f_Y(y)} = \frac{15(x^2y^2 - y^4)}{2f_Y(y)}.$$

A nevezőben szereplő sűrűségfüggvényt is számoljuk ki ($0 < y < 1$):

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx = \int_y^1 15xy^2dx = \frac{15}{2}(y^2 - y^4).$$

Így tehát a kérdéses eloszlásfüggvény ($0 \leq y \leq 1$ értékek esetén):

$$F_{X|Y}(x|y) = \frac{x^2y^2 - y^4}{y^2 - y^4} \text{ , ahol } y < x < 1.$$

A kérdéses sűrűségfüggvényt számoljuk ki kétféleképpen! Elsőként a 10. tétel értelmében ($0 < y < 1$ értékek esetén):

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{15xy^2}{\frac{15}{2}(y^2 - y^4)} = \frac{2x}{1 - y^2} \text{ , ahol } y < x < 1.$$

Másrészt számolhatjuk az eloszlásfüggvény deriválása által ($0 < y < 1$ értékek esetén):

$$f_{X|Y}(x|y) = \frac{d(F_{X|Y}(x,y))}{dx} = \frac{d\left(\frac{x^2y^2-y^4}{y^2-y^4}\right)}{dx} = \frac{2xy^2}{y^2 - y^4} = \frac{2x}{1 - y^2} \text{ , ahol } y < x < 1.$$

Látható, hogy valóban ugyanazt a sűrűségfüggvényt kaptuk a két különböző megközelítésből.

Emlékezzünk, hogy korábban már általánosítottuk a teljes valószínűség tételét arra az esetre, ha folytonos eloszlású értékre akartunk feltételezni. Azonban eddig nem beszéltünk arról az általánosítási lehetőségről, hogy valószínűség helyett kérdezhetnénk sűrűségfüggvényt is, amennyiben egy folytonos eloszlás viselkedését szeretnénk vizsgálni.

11. Tétel. *(Folytonos teljes valószínűség tétel sűrűségfüggvényre) Ha X és Y együttes eloszlása folytonos, akkor*

$$f_X(x) = \int_{-\infty}^{\infty} f_{X|Y}(x|t) f_Y(t)dt.$$

Bizonyítás: Két merőben eltérő módon bebizonyítjuk a tételt, hogy jobban rávilágítsunk az összefüggésekre.

1. *Tudjuk, hogy a perem sűrűségfüggvényt megkaphatjuk úgy, ha az együttes sűrűségfüggvényt integráljuk a teljes tartományon az összes felesleges változó szerint. Azaz:*

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy.$$

Használjuk fel a 10. Tételt:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X|Y}(x|y) f_Y(y) dy.$$

2. *Írjuk fel a folytonos teljes valószínűség tételt az $A = \{X < x\}$ eseményre és az Y folytonos valószínűségi változóra:*

$$F_X(x) = P(A) = \int_{-\infty}^{\infty} P(A|Y = t) f_Y(t) dt = \int_{-\infty}^{\infty} F_{X|Y}(x|t) f_Y(t) dt.$$

Ezt deriválva x szerint megkapjuk a sűrűségfüggvényt:

$$f_X(x) = \int_{-\infty}^{\infty} \frac{d(F_{X|Y}(x|t))}{dx} f_Y(t) dt = \int_{-\infty}^{\infty} f_{X|Y}(x|t) f_Y(t) dt.$$

Az előző példában szereplő X és Y valószínűségi változókra felírva a 11. tételt (ahol $0 < x < 1$):

$$f_X(x) = \int_{-\infty}^{\infty} f_{X|Y}(x|t) f_Y(t) dt = \int_0^x \frac{2x}{1-t^2} \frac{15}{2} (t^2 - t^4) dt =$$

$$= \int_0^x 15xt^2 dt = [5xt^3]_0^x = 5x^4.$$

Vegyük észre, hogy a második sorban pontosan az együttes sűrűségfüggvényt integráljuk a második változó szerint a teljes tartományon. Azaz amit számolunk megegyezik azzal, mint ahogy a korábbi ismereteink alapján számolnánk a vetületi sűrűségfüggvényt az együttes sűrűségfüggvényből.

Az eredeti teljes valószínűség tétel legfőbb alkalmazása a Bayes-tétel volt, aminek most szintén megfogalmazhatjuk a folytonos alakját:

12. Tétel. *(Folytonos Bayes-tétel) Amennyiben X és Y együttesen folytonos eloszlású valószínűségi változók, akkor*

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x) f_X(x)}{\int_{-\infty}^{\infty} f_{Y|X}(y|t) f_X(t) dt}$$

Bizonyítás: A 10. Tételből egyszerűen következik, hogy:

$$f_{X|Y}(x|y) f_Y(y) = f_{X,Y}(x, y) = f_{Y|X}(y|x) f_X(x),$$

amit átrendezve

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x) f_X(x)}{f_Y(y)}.$$

A 11. Tétel értelmében a nevező átírható pontosan az elvárt formára.

Az előző példákban szereplő X és Y valószínűségi változókra írjuk fel a 12. tételt, de fordított X és Y szerepekkel (mert $f_{X|Y}(x|y)$-t már ismerjük, de $f_{Y|X}(y|x)$-t még nem). Ekkor:

$$f_{Y|X}(y|x) = \frac{f_{X|Y}(x|y) f_Y(y)}{\int_{-\infty}^{\infty} f_{X|Y}(x|t) f_Y(t) dt}$$

Már kiszámoltuk a nevezőben szereplő integrált az előző példában, azt csak behelyettesítjük, így ahol $0 < x < 1$:

$$f_{Y|X}(y|x) = \frac{\frac{2x}{1-y^2} \frac{15}{2}(y^2 - y^4)}{5x^4} = \frac{3y^2}{x^3} \text{ , ahol } 0 < y < x.$$

A feltételes eloszlás esetén is fontos a várható érték vizsgálata:

13. Definíció. *Az $E(X|Y = y)$ feltételes várható érték az alábbi módon definiálható:*

1. *Ha Y diszkrét eloszlású*
$$E(X|Y = y) = \sum_{x \in R_X} x P(X = x|Y = y).$$
2. *Ha Y folytonos eloszlású*
$$E(X|Y = y) = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx.$$

14. Megjegyzés. *A feltételes várható érték egy függvény: $R_Y \to \mathbf{R}$; $y \mapsto E(X|Y = y)$. Ezt szokásosan regressziós függvénynek hívjuk. Az értelmezési tartománya Y értelmezési tartományával egyezik meg.*

Nézzük meg az előző példákban szereplő X és Y valószínűségi változók regressziós függvényeit:

$$E(X|Y = y) = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx = \int_y^1 x \frac{2x}{1 - y^2} dx =$$

$$= \left[\frac{2x^3}{3(1 - y^2)} \right]_y^1 = \frac{2 - 2y^3}{3(1 - y^2)} = \frac{2(1 + y + y^2)}{3(1 + y)} \text{ , ahol } 0 < y < 1.$$

$$E(Y|X=x) = \int_{-\infty}^{\infty} y f_{Y|X}(y|x)dy = \int_0^x y\frac{3y^2}{x^3}dy =$$

$$= \left[\frac{3y^4}{4x^3}\right]_0^x = \frac{3}{4}x \text{ , } ahol\ 0 < x < 1.$$

A feltételes várható érték tehát R_Y minden eleméhez értéket rendel. Azonban az $Y = y$ események $y \in R_Y$ esetén páronként kizáró események, amiknek az összege a biztos esemény. Tehát valójában a mögöttes objektum egy valószínűségi változó (hiszen bármely elemi eseményhez egyértelmű értéket rendel, és az egyes értékeihez tartozó események megfigyelhetőek).

15. Definíció. *Az X-nek az Y-ra vonatkozó* **regresszió***ja az a valószínűségi változó, ami egy elemi eseményhez, ami esetén az Y az y értéket veszi fel, az $E(X|Y=y)$ értéket veszi fel. Jelölése $E(X|Y)$.*

Vegyük észre, hogy ha az Y valószínűségi változót a regressziós függvénybe helyettesítjük, akkor pontosan a regressziót kapjuk.

16. Tulajdonság(ok). *A várható érték tulajdonságaiból következnek a regresszióra az alábbiak (X, Y, Z valószínűségi változók, a, b valós számok, $g(x)$ egy függvény):*

1. $E(aX + b|Y) = aE(X|Y) + b$,
2. $E(X + Z|Y) = E(X|Y) + E(Z|Y)$,
3. $E(E(X|Y)) = EX$,
4. $E(X|I_\Omega) = EX$ *(ahol I_Ω a biztos esemény indikátora),*
5. $E(X|X) = X$,
6. $E(g(X)|X) = g(X)$,
7. $E(g(Y) \cdot X|Y) = g(Y)E(X|Y)$,
8. *ha X és Y függetlenek, akkor* $E(X|Y) = EX$.

Emlékezzünk, hogy a várható érték minimalizálta a négyzetes különbségek átlagát. Ezzel analóg módon a regresszióra is kimondhatunk egy hasonló tételt:

17. Tétel. *Bármely $g : \mathbf{R} \to \mathbf{R}$ függvényre igaz, hogy*

$$E((X - g(Y))^2) \geq E((X - E(X|Y))^2).$$

Azaz az Y valószínűségi változó transzformációi közül pontosan az X-nek az Y-ra vonatkozó regressziója lesz várhatóan négyzetesen legközelebb X-hez.

Máshogy megfogalmazva, ha X viselkedését szeretnénk Y-nak egy transzformáltjával közelíteni, akkor a regresszió a legmegfelelőbb választás.

Nézzük meg az előző példákban szereplő X és Y valószínűségi változók regresszióit:

$$E(X|Y) = \frac{2(1 + Y + Y^2)}{3(1 + Y)}$$

$$E(Y|X) = \frac{3}{4}X.$$

Figyeljük meg, hogy a két regresszió nem egymás reciproka. Tehát míg Y-t X-ből $\frac{3}{4}X$-ként tudjuk legjobban megközelíteni, addig X-re nem a $\frac{4}{3}Y$ a legjobb lehetőségünk.

18. Következmény.

$$E((X - E(X|Y))^2) \leq \sigma^2 X$$

Bizonyítás: A 17. Tétel szerint a bal oldal bármilyen g függvény esetén felülről becsülhető $E((X - g(Y))^2)$-al. Ha egyszerűen a $g(Y) = EX$ konstans függvényt használjuk a becsléshez, akkor az $E((X - EX)^2)$ felső becslést kapjuk, ami pontosan $\sigma^2 X$.

A regresszió egy speciális esete, ha nem engedünk meg bármilyen $g(Y)$ függvényt, csak a lineáris transzformációkat:

19. Definíció. *X-nek az Y-ra vonatkozó* **lineáris regressziója** *az az $aY + b$ valószínűségi változó (ahol $a, b \in \mathbf{R}$), amire $E((X - (aY + b))^2)$ értéke minimális. Tehát $\forall a', b' \in \mathbf{R}$:*

$$E((X - (aY + b))^2) \leq E((X - (a'Y + b'))^2)$$

20. Tétel. *Ha $a = R(X, Y)\frac{\sigma X}{\sigma Y}$ és $b = EX - R(X, Y)\frac{\sigma X}{\sigma Y}EY$, akkor $aY + b$ az X-nek az Y-ra vonatkozó lineáris regressziója.*

Vizsgáljuk meg az előző példákban szereplő X és Y valószínűségi változók $aY+b$ és $cX+d$ lineáris regresszióit! Ehhez szükségünk lesz a várható értékekre, szórásokra, kovarianciára és korrelációra.

$$EX = \int_0^1 5x \cdot x^4 = \frac{5}{6}$$

$$E(X^2) = \int_0^1 5x^2 \cdot x^4 = \frac{5}{7}$$

$$\sigma^2 X = \frac{5}{7} - \frac{25}{36} = \frac{5}{252}$$

$$EY = \int_0^1 \frac{15}{2} y \cdot (y^2 - y^4) = \frac{15}{2}\left[\frac{y^4}{4} - \frac{y^6}{6}\right]_0^1 = \frac{5}{8}$$

$$E(Y^2) = \int_0^1 \frac{15}{2} y^2 \cdot (y^2 - y^4) = \frac{15}{2}\left[\frac{y^5}{5} - \frac{y^7}{7}\right]_0^1 = \frac{5}{7}$$

$$\sigma^2 Y = \frac{5}{7} - \frac{25}{64} = \frac{145}{448}$$

$$E(XY) = \int_0^1 \int_0^x xy \cdot 15xy^2 dy dx = \int_0^1 \left[\frac{15}{4} x^2 y^4\right]_0^x dx =$$

$$= \int_0^1 \frac{15}{4} x^6 dx = \left[\frac{15}{28} x^7\right]_0^1 = \frac{15}{28}$$

$$cov(X,Y) = \frac{15}{28} - \frac{5}{6} \cdot \frac{5}{8} = \frac{180-175}{336} = \frac{5}{336}$$

$$a = \frac{cov(X,Y)}{\sigma^2 Y} = \frac{\frac{5}{336}}{\frac{145}{448}} = \frac{4}{87}$$

$$b = EX - aEY = \frac{5}{6} - \frac{4}{87} \cdot \frac{5}{8} = \frac{580-20}{696} = \frac{70}{87}$$

$$c = \frac{cov(X,Y)}{\sigma^2 X} = \frac{\frac{5}{336}}{\frac{5}{252}} = \frac{3}{4}$$

$$d = EY - cEX = \frac{5}{8} - \frac{3}{4} \cdot \frac{5}{6} = 0$$

Tehát X-nek az Y-ra vonatkozó lineáris regressziója $\frac{4}{87}Y + \frac{70}{87}$, Y-nak az X-re vonatkozó lineáris regressziója pedig $\frac{3}{4}X$. Ez utóbbi persze nem lep meg minket, hiszen ez volt az $E(Y|X)$, ami azt jelenti, hogy a regresszió lineáris volt, így persze a lineáris regresszió megegyezik a regresszióval.

A lineáris regresszió számítása esetén az alábbi egyszerű észrevétel segíthet önmagunk ellenőrzésében:

21. Következmény. *X-nek az Y-ra vonatkozó lineáris regressziójának várható értéke EX.*

Bizonyítás: Írjuk fel a kérdéses várható értéket:

$$E(aY + EX - aEY) = aEY + EX - aEY = EX.$$

Az előző példákban szereplő lineáris regressziókat ellenőrizzük le:

$$E\left(\frac{4}{87}Y + \frac{70}{87}\right) = \frac{4}{87}\cdot\frac{5}{8} + \frac{70}{87} = \frac{20+560}{3\cdot 29\cdot 8} = \frac{29\cdot 5\cdot 4}{3\cdot 29\cdot 8} = \frac{5}{6} = EX.$$

$$E\left(\frac{3}{4}X\right) = \frac{3}{4}\cdot\frac{5}{6} = \frac{5}{8} = EY.$$

Látható, hogy tényleg teljesül a tulajdonság. Ha már ezeket ellenőriztük, ellenőrizhetjük X-nek az Y-ra vonatkozó regresszióját is, hiszen a 16. pontban említett tulajdonságok között a 3. szerint a várható értéke ennek is EX kell legyen:

$$E(E(X|Y)) = E\left(\frac{2(1+Y+Y^2)}{3(1+Y)}\right) = \frac{2}{3} + \frac{2}{3}E\left(\frac{Y^2}{1+Y}\right) =$$

$$= \frac{2}{3} + \frac{2}{3}\int_0^1 \frac{y^2}{1+y}\frac{15}{2}(y^2 - y^4)dy = \frac{2}{3} + 5\int_0^1 \frac{y^2\cdot y^2\cdot(1-y)\cdot(1+y)}{1+y}dy =$$

$$= \frac{2}{3} + 5\int_0^1 (y^4 - y^5)dy = \frac{2}{3} + 5\left[\frac{y^5}{5} - \frac{y^6}{6}\right]_0^1 = \frac{2}{3} + \frac{1}{6} = \frac{5}{6} = EX.$$

Tehát valóban teljesül ez a tulajdonság a regresszióra az esetünkben.

A 16. pontban szereplő harmadik tulajdonságot ($EX = E(E(X|Y))$) ellenőrzésképpen használtuk eddig, de néha ennek a fajta megfordított iránynak a segítségével számolhatunk várható értéket:

22. Következmény. *(Teljes valószínűség tétel várható értékre)*
Ha Y diszkrét, akkor $EX = \sum_{y\in R_Y} E(X|Y=y)P(Y=y)$.
Ha Y folytonos, akkor $EX = \int_{-\infty}^{\infty} E(X|Y=y)f_Y(y)dy$.

A bizonyítása teljesen analóg a valószínűségre vonatkozó teljes valószínűség tételével. Használjuk fel a korábbi képletet transzformált valószínűségi változó várható értékére a $g(y) = E(X|Y=y)$ behelyettesítésével:
Ha Y diszkrét, akkor $E(g(Y)) = \sum_{y\in R_Y} g(y)\cdot P(Y=y)$, így

$$EX = E(E(X|Y=y)) = \sum_{y\in R_Y} E(X|Y=y)\cdot P(Y=y).$$

Ha Y folytonos, akkor $E(g(Y)) = \int_{-\infty}^{\infty} g(y) \cdot f_Y(y)dy$, így

$$EX = E(E(X|Y=y)) = \int_{-\infty}^{\infty} E(X|Y=y) \cdot f_Y(y)dy.$$

Mindkét esetben a 16. pontban szereplő harmadik tulajdonságot használtuk ki az első egyenlőségben.

Egy bányában eltévedt egy ember, aki egy csomópontban áll, ahonnan három irányba mehet tovább. Az első irányt választva kijut a bányából egy óra alatt, a második irányt választva visszaérkezik ugyanide 2 óra alatt, a harmadik irányt választva pedig 3 óra alatt ér vissza ugyanide. Feltéve, hogy amikor visszaérkezik, nem jön rá, hogy ugyanott áll, és minden esetben, amikor választania kell, akkor egyenlő valószínűséggel választja bármelyik opciót, várhatóan mennyi idő múlva jut ki a bányából?
Legyen X a kijutás idejének valószínűségi változója, Y pedig az első döntésének az iránya ($Y = i$, ha az i-edik irányt választja). Ekkor felírva a 22. következményt a diszkrét esetben:

$$EX = \sum_{y \in R_Y} E(X|Y=y)P(Y=y) = \frac{1}{3}\sum_{y=1}^{3} E(X|Y=y)$$

Mivel ha visszaér a kiindulási pontra, akkor a hátralevő idő várható értéke megegyezik az eredeti várható értékkel, így ez tovább alakítható:

$$EX = \frac{1}{3}(1 + 2 + EX + 3 + EX) = 2 + \frac{2}{3}EX,$$

amiből azonnal következik, hogy $EX = 6$.

3. Kétdimenziós normális eloszlás, polinomiális eloszlás. A függetlenség és korrelálatlanság kapcsolata normális esetben. A regresszió normális esetben lineáris. A polinomiális eloszlás vetületei binomiálisak.

Ebben a fejezetben két nevezetes többdimenziós eloszlást vizsgálunk meg.

3.1. Kétdimenziós normális eloszlás

23. Definíció. *Azt mondjuk, hogy X és Y együttes eloszlása* **(kétdimenziós) normális**, *ha léteznek $\mu_1, \mu_2 \in \mathbf{R}$, $\rho \in (-1, 1)$ és $\sigma_1, \sigma_2 \in \mathbf{R}^+$ konstansok, hogy* *az együttes sűrűségfüggvényük*

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_1)^2}{\sigma_1^2}+\frac{(y-\mu_2)^2}{\sigma_2^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}\right)}.$$

Jelölése $(X,Y) \in N(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)$.

A kétdimenziós normális eloszlás vetületei normális eloszlásúak:

24. Következmény. *Ha* $(X,Y) \in N(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)$, *akkor* $X \in N(\mu_1,\sigma_1)$, $Y \in N(\mu_2,\sigma_2)$ *és* $R(X,Y) = \rho$.

Hogy miként állíthatunk elő ilyen eloszlást, vagy milyen motívum válthatja ki a valóságban, arról szól a következő tétel:

25. Tétel. *Akkor és csak akkor lesz* (X,Y) *eloszlása kétdimenziós normális, ha kifejezhetőek két független* U,V *normális eloszlású valószínűségi változóból* $a,b,c,d \in \mathbf{R}$ *konstansokkal az alábbi formában:*

$$X = aU + bV \quad \textit{és} \quad Y = cU + dV.$$

A Miku és a Lás légiposta a feladott csomagokat Lappföldről a térfogatuk és súlyuk alapján szállítja Budapestre. Ezeknek a különleges csomagoknak a térfogata $U \in N(10,2)$ *eloszlású (literben), míg a súlya* $V \in N(4,\sqrt{2})$ *eloszlású (kilogrammban), és a kettő egymástól független. A Miku díjszabása* $\frac{1}{2}$ *euró literenként és* 2 *kilogrammonként, míg a Lás díjszabása* 1 *euró literenként és* 0 *kilogrammonként (ünnepi akcióban). Mi lesz egy véletlenszerűen választott csomag esetén a két posta költségének együttes eloszlása?*
A 25. tétel értelmében az együttes eloszlás kétdimenziós normális lesz, hiszen U *és* V *független normálisok, a Miku ára* $X = \frac{1}{2}U + 2V$, *és a Lás ára* $Y = U + 0V$. *Tehát* $(X,Y) \in N(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)$, *ahol:*

$$\mu_1 = EX = E\left(\frac{1}{2}U + 2V\right) = 13 \quad , \quad \mu_2 = EY = E(U) = 10,$$

$$\sigma_1 = \sigma\left(\frac{1}{2}U + 2V\right) = \sqrt{\frac{1}{4}\sigma^2(U) + 4\sigma^2(V)} = \sqrt{\frac{1}{4}\cdot 4 + 4\cdot 2} = 3,$$

$$\sigma_2 = \sigma(U) = 2,$$

$$\rho = R\left(\frac{1}{2}U + 2V, U\right) = \frac{E((\frac{1}{2}U + 2V)U) - E(\frac{1}{2}U + 2V)EU}{\sigma_1\sigma_2} =$$

$$= \frac{\frac{1}{2}E(U^2) + 2EUEV - \frac{1}{2}(E(U))^2 - 2EUEV}{3\cdot 2} = \frac{\sigma^2(U)}{12} = \frac{1}{3}.$$

26. Tétel. *$(X,Y) \in N(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)$ esetén X és Y akkor és csak akkor függetlenek, ha $\rho = 0$.*

Ha függetlenek, akkor természetesen a korreláció nulla. Ha pedig $\rho = 0$, akkor a képlet szerinti sűrűségfüggvényre igaz, hogy:

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_1\sigma_2} e^{\frac{-1}{2}\left(\frac{(x-\mu_1)^2}{\sigma_1^2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right)} =$$

$$= \frac{1}{\sigma_1\sqrt{2\pi}} e^{\frac{-(x-\mu_1)^2}{2\sigma_1^2}} \frac{1}{\sigma_2\sqrt{2\pi}} e^{\frac{-(y-\mu_2)^2}{2\sigma_2^2}} = f_X(x) f_Y(y).$$

Azaz az együttes sűrűségfüggvény pontosan a vetületi sűrűségfüggvények szorzata, tehát a két valószínűségi változó független.

A kétdimenziós normális eloszlás vetületei közötti regresszió lineáris:

27. Tétel. *Ha $(X,Y) \in N(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)$, akkor*

$$E(X|Y) = \left(\frac{\sigma_1}{\sigma_2}\rho\right) Y + \left(\mu_1 - \frac{\sigma_1}{\sigma_2}\rho\mu_2\right).$$

Határozzuk meg az előző példában szereplő postázási költségek regresszióit!

$$E(X|Y) = \left(\frac{\sigma_1}{\sigma_2}\rho\right) Y + \left(\mu_1 - \frac{\sigma_1}{\sigma_2}\rho\mu_2\right) = \left(\frac{3}{2}\frac{1}{3}\right) Y + \left(13 - \frac{3}{2}\frac{1}{3}10\right) = \frac{1}{2}Y + 8,$$

$$E(Y|X) = \left(\frac{\sigma_2}{\sigma_1}\rho\right) X + \left(\mu_2 - \frac{\sigma_2}{\sigma_1}\rho\mu_1\right) = \left(\frac{2}{3}\frac{1}{3}\right) X + \left(10 - \frac{2}{3}\frac{1}{3}13\right) = \frac{2}{9}X + \frac{64}{9}.$$

Ahogy korábban is tettük, ellenőrizzük a regressziókat, hogy a várható értéke tényleg az elvárt-e:

$$E(E(X|Y)) = E\left(\frac{1}{2}Y + 8\right) = \frac{1}{2}10 + 8 = 13 = EX,$$

$$E(E(Y|X)) = E\left(\frac{2}{9}X + \frac{64}{9}\right) = \frac{2}{9}13 + \frac{64}{9} = 10 = EY.$$

3.2. Polinomiális eloszlás

A binomiális eloszlás kísérletét elképzelhetjük úgy, mintha egy urnából visszatevéssel húznánk golyókat, és a pirosak számát figyelnénk. Mi lenne a helyzet, ha nem csak egy színt szeretnénk megfigyelni, azaz például ha a húzott pirosak, zöldek és kékek számát is külön-külön szeretnénk megfigyelni?

28. Definíció. *Legyen egy kísérleten az* $A_1, A_2, ..., A_k$ *teljes eseményrendszer olyan, hogy* $p_i = P(A_i) > 0$*. Elvégezzük a kísérletet* n*-szer (teljesen függetlenül), és az* X_i *valószínűségi változó azon kísérletek száma, amelyekben* A_i *bekövetkezett. Ekkor azt mondjuk, hogy* $(X_1, X_2, ..., X_k)$ **polinomiális eloszlás***ú. Jelölése*

$$(X_1, X_2, ..., X_k) \in Pol(n, p_1, p_2, ..., p_k).$$

29. Következmény. *Ha* $(X_1, X_2, ..., X_k) \in Pol(n, p_1, p_2, ..., p_k)$*, akkor* $x_1, x_2, ..., x_k$ *nemnegatív egészek esetén, ahol* $\sum_{i=1}^{k} x_i = n$

$$P(X_1 = x_1, X_2 = x_2, ..., X_k = x_k) = \frac{n!}{x_1!\, x_2! \cdots x_k!} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}.$$

30. Tulajdonság(ok). *Ha* $(X_1, X_2, ..., X_k) \in Pol(n, p_1, p_2, ..., p_k)$*, akkor*

1. $\sum_{i=1}^{k} X_i = n$,
2. $X_i \in Bin(n, p_i)$, *azaz az egydimenziós vetületi eloszlások binomiálisak,*
3. $X_{i_1} + X_{i_2} \in Bin(n, p_{i_1} + p_{i_2})$,
4. $(X_1 + X_2, X_3, X_4, ..., X_k) \in Pol(n, p_1 + p_2, p_3, p_4, ..., p_k)$,
5. $E(X_1, X_2, ..., X_k) = (np_1, np_2, ..., np_k)$,
6. $cov(X_{i_1}, X_{i_2}) = -np_{i_1} p_{i_2}$,
7. $R(X_{i_1}, X_{i_2}) = \frac{-\sqrt{p_{i_1}}\sqrt{p_{i_2}}}{\sqrt{1-p_{i_1}}\sqrt{1-p_{i_2}}}$.

4. Nevezetes valószínűségszámítási paradoxonok

A valószínűségszámítási paradoxonok kivétel nélkül arra építenek, hogy a naiv elgondolásunk milyen hibákat tud okozni a gondolatmenetünkben. Sokszor egy-egy egyszerű problémafelvetésre kézenfekvőnek tűnik a megoldás, azonban ezek mögött hibák húzódhatnak, ha nem precízen implementáljuk a valószínűségszámítást. A jegyzetben idáig elérve, az olvasónak már megvan minden alapja ahhoz, hogy megértse az alábbi paradoxonok érvelési hibáját, így fel tudjuk oldani azokat. Ez a fajta készség fontos a való életben, mert ezekkel ekvivalens hibákat nap mint nap vétenek az emberek.

Fontos, hogy nem elég, ha egy problémát meg tudunk oldani helyesen, fel kell ismernünk a hibát a rossz megoldásban. Ellenkező esetben több megoldás közül nem lennénk képesek eldönteni, hogy melyik a helyes.

4.1. Monty Hall paradoxon

Egy televíziós műsorban (amit sokáig Monty Hall vezetett) a játékos előtt van három zárt ajtó. Kettő ajtó mögött egy-egy kecske van, a harmadik mögött pedig egy vadonatúj autó (ami többet ér, mint egy kecske). A játékos választ egy ajtót, ezután a műsorvezető kinyit egy másikat, ami mögött kecske van. Ezután a játékos dönthet úgy, hogy az eredeti döntésénél marad, és elviszi a választott ajtó mögötti dolgot, de dönthet úgy is, hogy a másik még zárt ajtó mögötti nyereményt kéri. Megéri cserélni?

Rossz érvelés: Nem nyerünk a cserén, hiszen klasszikus valószínűségi mezőből indultunk, amit nem ront el egy ajtó kinyitása (csak az értékkészletet szűkíti). Így a maradék két ajtó mögött $\frac{1}{2}$ valószínűséggel lehet az autó. Precízebben: Számozzuk az ajtókat a választásunk alapján: 1-es amit választottunk, 2-es amit kinyit a műsorvezető, 3-as az utolsó. Legyen A, B, C az az esemény, hogy az autó rendre az 1-es, 2-es, 3-as ajtó mögött van. Ekkor

$$P(A|\overline{B}) = \frac{1}{2} \quad \text{és} \quad P(C|\overline{B}) = \frac{1}{2}$$

Nem igaz, hogy egy ajtó kinyitása nem változtatja meg a valószínűségeket. Emlékezzünk, hogy egy eseményre feltételezve akkor marad ugyanaz a valószínűség, hogyha a két esemény független volt egymástól. Az ajtók fenti számozása sem lehetséges, hiszen a műsorvezető biztosan nem kecskés ajtót nyit ki, így a számozás már alapból csak a feltételes „világban” él.

Mielőtt megnéznénk a jó megoldást, történeti áttekintés mellett megemlítjük a félreértés egy lehetséges okát:

A televízióban a 60-as évektől futott a műsor, aminek matematikai problémáját 1975-ben vetette fel, és oldotta meg Steve Selvin. A matematikai problémát akkor kapták fel igazán, amikor 1990-ben egy újságban Marilyn vos Savant-nak (őt a világ akkori legokosabb embereként tartották számon) címezték kérdésként. Marilyn helyes választ adott, de ezzel olyan sokan nem értettek egyet, hogy a felháborodott olvasók nagyságrendileg tízezer levelet írtak, amiből közel ezret PhD fokozattal rendelkező emberek küldtek. Még Erdős Pál sem hitte el (1995-ben), hogy a váltás a megfelelő, amikor Vázsonyi Endre felvetette neki a problémát (aki szintén csak szimulációkkal tudta magát meggyőzni). Néhány emberben a félreértés abból eredhet, hogy nem teljesen tiszta a játékvezető szerepe (feltételezem, hogy Erdősnél is ez lehetett a probléma). Ha a játékvezető nem ismerné, hogy melyik ajtó mögött mi van, akkor tényleg nem változna meg a valószínűség. Azonban ebben az esetben lehetséges volna, hogy a játékvezető az autót fedi fel, amit a show szempontjából el akartak kerülni. Tehát tegyük fel, hogy a játékvezető előre megfontolt szándékkal mindenképp kecskés ajtót nyit ki.

Váltás nélkül persze egyértelmű, hogy $\frac{1}{3}$ a valószínűsége annak, hogy megnyerjük az autót. Azonban az ajtó megváltoztatásával mi lesz a valószínűsége annak, hogy autót találunk? Használjunk teljes valószínűség tételt ($AL = \{$autó lesz az új ajtó mögött$\}$; $AV = \{$autó volt az eredetileg választott ajtó mögött$\}$; $KV = \{$kecske volt az eredetileg választott ajtó mögött$\}$):

$$P(AL) = P(AL|AV)P(AV) + P(AL|KV)P(KV) = 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3}.$$

Azaz látható, hogy kétszer akkora esélyünk van autót nyerni, ha a váltás mellett döntünk. Ha valaki még mindig nem lenne meggyőzve, érdemes elgondolni a helyzetet 100 ajtóval és 99 kecskével. 1% lenne az esélye annak, hogy elsőre jó ajtót választunk. Tegyük fel, hogy a játékvezető kinyit 98 ajtót (és csak kecskéset nyithat). Ekkor az az esemény, hogy a másik ajtó mögött van az autó ekvivalens azzal, hogy az első választásunk nem az autót rejtő ajtó volt. Tehát 99% az esélye, hogy a váltással nyerjük meg az autót.

31. Megjegyzés. *A Monty Hall paradoxonnak ekvivalens megfogalmazása a három rab paradoxon. Ott a rabok felelnek meg az ajtóknak, és egyikőjüket kivégzik, ez felel meg az autónak.*

4.2. Fiú vagy lány paradoxon

Andrásnak és Bíbornak is két gyereke van. András idősebbik gyereke lány, és Bíbornak is van lánya. Mi annak a P_A valószínűsége, hogy Andrásnak két lánya van, illetve annak a P_B valószínűsége, hogy Bíbornak két lánya van?

Rossz érvelés: Mindkét esetben egy gyereknek ismerjük a nemét, a másik pedig egyforma valószínűséggel lesz lány vagy fiú. Így a két valószínűség megegyezik, az értékük $\frac{1}{2}$.

Hibás megfogalmazás, hogy az egyik gyerek nemét ismerjük. András esetén valóban $\frac{1}{2}$ lesz a kérdéses valószínűség, mert a fiatalabbik gyerekről nem volt információnk. Ezzel ellenben Bíbor esetén az ismeretünk, hogy van lány gyermeke, egyformán ad információt a két gyerekről. Bíbor két gyereke alapján négy eset lehetséges (lány-fiú, lány-lány, fiú-lány, fiú-fiú), mind a négy egyforma valószínűségű (klasszikus valószínűségi mező). A kérdéses feltételes valószínűség így $\frac{1}{3}$, hiszen egy eset van a metszetben, és 3 eset van a feltételben.

4.3. Bertrand dobozai paradoxon

Tegyük fel, hogy három dobozunk van. Az egyikben két arany érme van, a másikban két ezüst, a harmadikban pedig egy arany és egy ezüst. Véletlenszerűen választottam egy dobozt, és kihúztam belőle egy érmét. Ha az érme arany, akkor mi a valószínűsége annak, hogy a dobozban levő másik érme is arany?

Rossz érvelés: Biztos, hogy nem a második doboz van nálam. Ha az első van nálam, akkor arany a másik érme, ha pedig a második, akkor ezüst, így $\frac{1}{2}$ a kérdéses valószínűség. Teljes valószínűség tétellel

$$P(A) = P(A|1)P(1) + P(A|3)P(3) = 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} = \frac{1}{2}.$$

A hiba az érvelésben az, hogy nem egyforma valószínűségű a két esemény, hogy az 1-es vagy a 3-as doboz van nálam. Az, hogy aranyat húztam, egy olyan esemény, ami megváltoztatja annak a valószínűségét, hogy az 1-es dobozt választottam-e. A helyes teljes valószínűség tétel a következő lenne (ahol F az az esemény, hogy az elsőnek húzott érme arany):

$$P(A|F) = P(A|1,F)P(1|F) + P(A|3,F)P(3|F) = 1 \cdot \frac{2}{3} + 0 \cdot \frac{1}{3} = \frac{2}{3}.$$

Ez azért igaz, mert ha az első húzásom arany volt, akkor $\frac{2}{3}$ valószínűséggel az 1-es doboz van a kezemben.

4.4. Nem tranzitív kocka paradoxon

Tekintsünk olyan dobókockákat, amiknek az oldalain pozitív egészek szerepelhetnek. Mondjuk azt, hogy egy ilyen kocka jobb egy másiknál, ha több, mint $\frac{1}{2}$ valószínűséggel nagyobbat dobunk vele, mint a másikkal.

Rossz érvelés: Ha egy A kocka jobb egy B kockánál, és B jobb egy C kockánál, akkor A jobb, mint C. Azaz a jobb reláció tranzitív.

$$P(A > B) > \frac{1}{2} \text{ és } P(B > C) > \frac{1}{2} \implies P(A > C) > \frac{1}{2}$$

A fenti következtetés egyszerűen nem igaz. Még ha függetlenek lennének az $A > B$ és a $B > C$ események, akkor is csak $\frac{1}{4}$ alsó korlát adódna az $A > C$ valószínűségre a metszetükben.

Létezik olyan kocka hármas A, B és C, hogy bármelyik kockánál van jobb a másik kettő között. Legyen a három kockánk valószínűségi változójának értékkészlete $R_A = \{2,2,4,4,9,9\}, R_B = \{1,1,6,6,8,8\}, R_C = \{3,3,5,5,7,7\}$. Ekkor $P(A > B) = \frac{5}{9} > \frac{1}{2}$ és $P(B > C) = \frac{5}{9} > \frac{1}{2}$, de $P(C > A) = \frac{5}{9} > \frac{1}{2}$. Ezekkel a kockákkal elképzelhető egy olyan játék, aminél egy másik paradoxonhoz juthatunk:

András és Bíbor választani fog egy-egy kockát, amivel egymás után sokszor fognak egyszerre dobni. Minden kört az nyer, aki a nagyobb értéket dobta. András választ először egy kockát a fenti A, B és C közül, Bíbor pedig a maradék kettőből választ egyet.

Rossz érvelés: András jár jobban, hiszen ő választhat először, Bíbor csak a maradékból vehet.

A fent írt valószínűségek szerint Bíbor a maradékból tud olyan kockát választani, amivel $\frac{5}{9}$ valószínűséggel fog nyerni minden körben. Tehát akárhogyan is választ András, Bíbor jár jobban.

4.5. Két boríték paradoxon

András pénzt tesz két borítékba, az elsőbe α-t a másikba 2α-t. Bíbornak (aki nem ismeri α értékét) felajánlja, hogy válasszon egy borítékot, és nézzen bele. Ezután Bíbort döntés elé állítja: megtarthatja a kinyitott boríték tartalmát, vagy dönthet úgy, hogy a másik boríték tartalmát kéri. Melyik döntés lesz jobb?

Rossz érvelés: Tudjuk, hogy $\frac{1}{2}$ valószínűséggel húztuk a kisebbiket, és ugyanennyivel a nagyobbikat. Ha X van a húzott borítékban, akkor a másik borítékban levő pénzmennyiség várható értéke:

$$E(Y|X=x) = \frac{1}{2} \cdot \frac{x}{2} + \frac{1}{2}2x = \frac{5}{4}x > x.$$

Így tehát jobban megéri a másikat választani.

Persze mivel ez a döntés független volt attól, hogy mit látunk a kinyitott borítékban, rögtön érezzük, hogy nem lehet igaz. Ha úgy számolnánk, hogy az elsőnek húzott boríték várható értéke $EX = \frac{1}{2}\alpha + \frac{1}{2}2\alpha = \frac{3}{2}\alpha$, akkor látható, hogy ugyanez igaz a második borítékra is, így a váltás se nem ront, se nem javít a nyereményünk várható értékén. Mi a hiba a korábbi EY számításában?

A hiba az, hogy X valószínűségi változó értékét fixálva valójában nem marad két lehetőség. Ha $x = \alpha$, akkor 1 valószínűséggel 2α van a másikban, és ha $x = 2\alpha$, akkor 1 valószínűséggel α van a másikban. A várható érték a fenti képlethez hasonlóan úgy lenne korrekt, hogy

$$EY = \frac{1}{2}E(\frac{X}{2}|X > Y) + \frac{1}{2}E(2X|X < Y) =$$

$$= \frac{1}{2}E(\frac{X}{2}|X = 2\alpha) + \frac{1}{2}E(2X|X = \alpha) = \frac{1}{2}\alpha + \frac{1}{2}2\alpha = \frac{3}{2}\alpha.$$

Így pedig látszik, hogy tényleg nem nyerünk azzal, ha a váltást választjuk.

32. Megjegyzés. *Figyeljük meg, hogy míg a Monty Hall paradoxonban megérte váltani, itt nem. A különbség abból származik, hogy a korábbinál a műsorvezető tudása miatt a kecskés ajtó kinyitása megváltoztatta a valószínűségeket, míg utóbbinál a kihúzott boríték tartalmának megnézése nem adott információt arról, hogy melyik borítékban van a nagyobb összeg.*

4.6. Nyakkendő paradoxon

Xavier és York nyakkendőt kaptak karácsonyra a feleségeiktől. Egy beszélgetésben azon tanakodnak, melyiküké az olcsóbb. Kieszelik azt a fogadást, miszerint megkérdezik a feleségeiket az áraikról, és akié a drágább, annak oda kell adnia a nyakkendőjét a másiknak. Feltételezhetjük, hogy a két nyakkendő ára - X és Y - azonos eloszlású, és együttes eloszlásuk szimmetrikus, azaz X és Y szerepe felcserélhető (ez a kritérium kicsit gyengébb feltevés, mint a függetlenség).

Rossz érvelés: Bármelyik férfi érvelhet úgy, hogy rossz esetben a saját nyakkendőjét veszíti el, azonban jó esetben egy annál értékesebbet nyer.

Pontosabban számolva, ha Xavier nyeresége A_X, akkor:

$$E(A_X|X<Y) = E(Y|X<Y) > EX.$$

Azaz mivel ha nyer, akkor várhatóan többet nyer, mint amennyit veszíteni tud, így megéri játszani.

Így mindkettőjük részéről pozitívnak tűnik a várható érték, ami persze lehetetlen, hiszen az összeg 0 kell legyen. A fenti érvelésben az a hiba, hogy valójában amikor veszít, akkor is többet veszít, mint EX. A két esetre vett teljes várható érték (felhasználva, hogy azonos eloszlásúak):

$$E(A_X) = E(Y|X<Y)\cdot P(X<Y) - E(X|X>Y)\cdot P(X>Y) = 0.$$

A várható érték számítását részletesen így fejthetnénk ki (ha az ár eloszlása folytonos eloszlású):

$$E(A_X) = \int_{-\infty}^{\infty}\int_{-\infty}^{y} y f_{X,Y}(x,y)dxdy - \int_{-\infty}^{\infty}\int_{y}^{\infty} x f_{X,Y}(x,y)dxdy =$$

$$= E(Y) - \int_{-\infty}^{\infty}\int_{y}^{\infty} y f_{X,Y}(x,y)dxdy - \int_{-\infty}^{\infty}\int_{-\infty}^{x} x f_{X,Y}(x,y)dydx =$$

$$= E(Y) - \int_{-\infty}^{\infty}\int_{y}^{\infty} y f_{X,Y}(x,y)dxdy - \int_{-\infty}^{\infty}\int_{-\infty}^{y} y f_{X,Y}(x,y)dxdy = E(Y) - E(Y) = 0.$$

33. Megjegyzés. *A két boríték paradoxon speciális esetként előállítható a nyakkendő paradoxonból, ha előre tudnánk, hogy a két nyakkendő ára α és 2α.*

34. Megjegyzés. *A két boríték paradoxon és a nyakkendő paradoxon is elképzelhető úgy, mint valuták közötti váltás, ez a verzió a Siegel paradoxon. Például tegyük fel, hogy 1-1 váltás lehetséges most az A és a B valuta között, de holnap valamelyik valuta kétszeresét fogja érni a másiknak (egyforma valószínűséggel). Ekkor mindkét valuta birtokosa érvelhet úgy (hibásan), hogy átváltva a pénzét, holnapra várhatóan $\frac{5}{4}$-szeresét éri majd az összeg annak a verziónak, mintha nem váltanánk át.*

4.7. Bertrand-féle húr paradoxon

Válasszuk ki egy egységsugarú kör egy húrját véletlenszerűen! Mi a valószínűsége annak, hogy a választott húr hossza legalább $\sqrt{3}$ (azaz a körbe írható egyenlő oldalú háromszög oldalhossza)?

Rossz érvelés: Bármilyen megközelítést alkalmazhatunk a probléma modellezésére:

1. lehet rögzített a választott húr iránya, és az irányra merőleges átmérőn egyenletesen helyezkedhet el a húr és az átmérő metszéspontja.
2. lehet rögzített a húr egyik végpontja, és a másikat választhatjuk egyenletesen a kör kerületén.
3. egyenletesen választhatunk egy pontot a kör belsejéből, és illesszük rá azt a húrt, aminek pont ő a középpontja (a kör középpontján kívül minden belső pontjára egyértelmű ez a húr, de a hossza még a középpontra is az).

Akármelyik modellből számolható a kérdéses valószínűség.

A hiba a fenti érvelésben, hogy a három fent definiált eloszlás különböző, és a kapott valószínűségek is eltérőek lesznek. Valójában a kérdésfelvetés pontatlan, mert sok különböző véletlen módon választhatunk húrt. Fontos megértenünk, hogy a véletlen szó önmagában nem specifikál eloszlást. Legtöbbször a véletlen alatt az egyenletes eloszlást értjük, de ez a példa jól mutatja, hogy fontos specifikálni a paramétert, hogy mire vonatkozik az egyenletesség.

Számoljuk ki a fenti modellek esetén a kérdéses valószínűséget:

1. Ha rögzített a választott húr iránya, és egyenletes a merőleges átmérőn a metszéspont: Képzeljük el azt a $\sqrt{3}$ oldalú szabályos háromszöget a körön, aminek az említett átmérő az AB oldalát felező magasságvonala. A kör középpontjának és az AB-nek a távolsága ekkor $\frac{1}{2}$. Ebből azonnal adódik, hogy a kedvező esetek a középponthoz $\frac{1}{2}$-nél közelebb helyezkednek el. Az átmérő hossza 2, a kedvező hossz 1, így a kérdéses valószínűség $\frac{1}{2}$.

2. Ha rögzített a húr egyik P végpontja, és a másik egyenletes a kör kerületén: Képzeljük el azt a $\sqrt{3}$ oldalú szabályos háromszöget a körön, aminek az egyik pontja P. A másik kettő pontja legyen Q és R. Ekkor a kör kerületén pontosan azok lesznek a kedvező esetek, amik Q és R között helyezkednek el. A három körív (PQ, QR és RP) egyforma hosszú, így a kérdéses valószínűség $\frac{1}{3}$.

3. Egyenletesen választva pontot a kör belsejéből a húr középpontjának: A kör középpontjának a $\sqrt{3}$ hosszú húroktól vett távolsága $\frac{1}{2}$. Ez azt jelenti, hogy pontosan akkor hosszabb $\sqrt{3}$-nál egy húr, ha közelebb van a középponthoz, mint $\frac{1}{2}$. Így a kérdéses valószínűség a két kör területének a hányadosából számolható: $\frac{\frac{1}{4}\pi}{\pi} = \frac{1}{4}$.

Látható, hogy az eloszlástól függően három különböző értéket kaptunk a valószínűségre.

4.8. Simpson-paradoxon

A tantárgy során nem vizsgáltuk a statisztika témakörét mélyebben, itt is csak annyira mélyülünk el a paradoxonban, hogy felhívjuk a figyelmet a lehetséges rossz elgondolásra. Tegyük fel, hogy András és Bíbor minden nap kapnak feladatokat (a feladatok száma változó, és nem feltétlen egyezik meg kettőjüknél). Azt tapasztaljuk, hogy András minden egyes napon jobban teljesít, mint Bíbor, azaz az elvégzett és a kapott feladatok aránya nála magasabb, mint Bíbornál. Következik-e ebből, hogy András összességében is jobban teljesített Bíbornál?

Rossz érvelés: Persze, hogy következik, hiszen ha minden nap jobb volt, akkor az összes nap együttesén (amire összegként, vagy átlagként tekinthetünk) is jobbnak kellett lennie.

Hibás az elgondolás, mert az összegzés igazából egy súlyozott összeg a háttérben, és a súlyozás megváltoztathatja az arányokat. Például ha az alábbiak lennének az adatok:
Első nap András 1 feladatot kapott, amit teljesített, Bíbor pedig 9-et kapott, amiből 8-at sikerült megoldani (persze $1 > \frac{8}{9}$).
Második nap András 9 feladatot kapott, amiből 1-et teljesített, Bíbor pedig 1-et kapott, de azt nem sikerült megoldani (persze $\frac{1}{9} > 0$).
Összességben figyelve azonban $\frac{2}{10} < \frac{8}{10}$, azaz Bíbor négyszer olyan jól teljesített (András 20%-os, Bíbor pedig 80%-os sikeraránnyal dolgozott).

A Simpson-paradoxon sok statisztikai hibás eredmény alapja. Ezt a fajta hibát angolul egyszerűen „bias”-nak is szokták hívni. Nézzünk erre egy megtörtént esetet:

Két vesekő kezelésére szolgáló módszer sikerességét akarták megvizsgálni (1986-ban). Ehhez 350 esetet gyűjtöttek mind a kettő kezelésre (figyeljünk oda, hogy már az esetek gyűjtésének a módja is torzíthat, például ha élő betegeket kérdezünk meg, akkor kimaradnak az olyanok, akik elhaláloztak a kezelés közben). Az első kezelés 273, míg a második 289 sikert hozott.

Rossz érvelés: Tehát a második jobb kezelés, ezután csak azt érdemes folytatni.

Az adatokat pontosabban megvizsgálva kiderült, hogy 87-szer alkalmazták kisebb vesekövek esetén az első kezelést, és 270-szer a másodikat. Így kiszámolható, hogy a kisebb kövekre 93%-os az első, míg 87%-os a második kezelés sikeraránya, továbbá nagyobb kövekre 73%-os az első, és 69%-os a második kezelés sikeraránya. Tehát valójában az első kezelés ezek alapján jobbnak mondható. Óvakodjunk az elhamarkodott kijelentésektől! Ha ezek után még kiderülne, hogy a vizsgált esetek között vannak sürgősségi műtétek, amik már belső vérzés beindulása után voltak kivitelezve, akkor ezeknek a száma is befolyásolhatja a performanciát (ez már nem része a valós tanulmánynak). Nézzük például az alábbi táblázatot:

Első kezelés				Második kezelés			
273/350 78%				289/350 83%			
Kis kő		Nagy kő		Kis kő		Nagy kő	
81/87 93%		192/263 73%		234/270 87%		55/80 69%	
Sürgős	Tervezett	Sürgős	Tervezett	Sürgős	Tervezett	Sürgős	Tervezett
81/85	0/2	192/200	0/63	200/200	34/70	50/50	5/30

Tehát ha pontosan ez a két befolyásoló tényező volt csak, ami különbséget tehet a vizsgált esetek között, akkor a második kezelés az, ami jobbnak mondható.

Azaz amikor össze akarunk hasonlítani két dolgot, akkor minden más körülménynek azonosnak kell lennie, hogy ezt a hibát elkerülhessük.

4.9. Schrödinger macskája

A statisztika után a fizika területén is érdemes egy valószínűségszámítási problémakört megimernünk. Schrödinger macskája egy gondolatkísérlet, amivel a szuperpozíció elvét szokták szemléltetni (és annak a makro világban való abszurditását). Meg fogjuk mutatni, hogy a valószínűségszámításban valójában egy nagyon egyszerű koncepció rejlik a kísérlet mögött. A matematikai elemzés rávilágít, hogy vagy a szemléltető példa hibás, vagy a szemléltetni kívánt szuperpozíció az, amit teljesen félreértelmeznek.

Tegyük fel, hogy bezárunk egy élő macskát egy dobozba, és egy nap múlva vesszük elő a dobozt ismét. Mondhatjuk, hogy a macska életben van?

Rossz érvelés: A macskáról egészen addig nem tudjuk, hogy életben van-e, amíg ki nem nyitjuk a dobozt. Tehát a macska a kinyitás előtt egyszerre élő és halott. Ezt nevezzük úgy, hogy a macska állapota az élő és a halott állapot szuperpozíciójában van.

Tényleg nem tudhatjuk biztosra, hogy a macska él-e, azonban hibás az állítás, hogy a macska állapota ezért egyszerre lenne élő és halott. A macska állapota egyértelmű, csak az a mi számunkra nem ismert. A mi ismeretünk egy X indikátor valószínűségi változónak tekinthető, ami 0, ha a macska halott, és 1, ha él.
Képzeljük el az egészet egy érmedobásként: az érmét feldobjuk, majd leérkezése pillanatában letakarjuk. Ekkor amíg fel nem fedjük, addig az érme se nem fej se nem írás a számunkra, de ez nem jelenti azt, hogy a kettő állapotot egyszerre birtokolja.

Idézve a Wikipédia megfelelő szócikkéből: „A kvantummechanikában szuperpozíciónak nevezik, amikor egy elemi részecske (vagy részecskékből álló rendszer) ún. kevert állapotban van (Megjegyzés: a magyar Wikipédián sajnos ez hibás, mert nem kell kevert állapotban lenni a szuperpozícióhoz, a két fogalom külön jelentéssel bír), azaz bizonyos tulajdonságait nem tudjuk egyértelműen. A részecske addig marad ebben, amíg valamilyen módon meg nem állapítjuk, hogy valójában hol és milyen állapotban van. A probléma ott kezdődik, hogy mérés (megfigyelés) hatására a részecske hullámfüggvénye összeomlik, és a részecske a lehetséges alap- vagy sajátállapotai egyikébe kerül, legalábbis minden általunk elvégezhető mérés azt mutatja, hogy a részecske egy bizonyos állapotban van.
Fontos megjegyezni, hogy a szuperpozíció (akár a hullámfüggvény) csakis abban az esetben omlik össze, ha mérést végzünk a rendszeren (vagy a rendszert valami külső hatás éri). A mérés eredményének ismerete azonban kényszerűen valamelyik állapotba taszítja az anyagot, ami Schrödingernél döglött vagy élő macskát eredményez, de sosem egyszerre a kettőt."

Az egész koncepció, hogy a mérés taszítja valamilyen állapotba a megfigyelt anyagot, a klasszikus világban lehetetlennek tűnik. Azonban a valószínűségi változók szintjén értékelve csak arról van szó, hogy a feltételes eloszlás nem egyezik meg az eredetivel. De ezt mi már ismerjük, hiszen $X|Y$ eloszlása természetesen nem azonos X eloszlásával. Azaz a mérés kimenetének ismeretében mást mondhatunk X-ről, de ettől még X nem változott meg. Pontosabban fogalmazva: X és $X|X = x$ eloszlása nem azonos, mert az $X = x$ feltevés az addig ismeretlen valószínűségi változót egy állapotába „taszítja".

5. Köszönet

A jegyzet készítése folyamán sok segítséget kaptam a valószínűségszámítás tárgy gyakorlatvezetőitől. Külön kiemelném Szabó Dánielt, aki a teljes anyagot tüze-

tesen átnézte, és a segítségével mind a matematikai, mind a nyelvi hibák számát nagymértékben csökkentettük.
Remélem, hogy a jegyzet eléri a célját, és megfelelő segítséget nyújt a tananyag elsajátításához az olvasóknak.