Lipécz György*

A HÁROM RAB ESETE

EGY VALÓSZÍNŰSÉG-SZÁMÍTÁSI PARADOXON TANULSÁGAI

Az alábbi paradoxon közismert, a valószínűség-számítási példatárak, logikai feladványgyűjtemények kedvelt darabja. A feladat első látásra nem tűnik bonyolultnak, de sokszor matematikában jártas emberek is belezavarodnak, és sokan a megoldást is kétkedve fogadják. Elemzésünkben ennek okait kutatjuk, bemutatva és több oldalról megvilágítva a megoldást és a megoldáshoz vezető út buktatóit.

A paradoxon

Három halálraítélt rab – A, B és C – szomorkodik a börtönben. Igaz, egyikük kegyelmet kapott, de nem tudják melyikük. Annyit tudnak: a kiválasztás véletlenszerű volt, az esély tehát egyenként 1/3. A börtönőr a faggatásra azt válaszolja: tilos neki elárulni, ki kapta a kegyelmet. Az A rab így szól a börtönőrhöz:

– Jó, nem azt kérdezem tőled, hogy én vagyok-e a szerencsés, aki kegyelmet kap. De legalább mutass rá a másik két társam közül az egyikre, aki biztos meghal. Ezzel nem árulsz el titkot, hiszen azt úgy is tudom, hogy a másik kettő közül legalább az egyiknek meg kell halnia.

Az őr némi tépelődés után enged a kérésnek, rámutat A egyik társára, C-re. A elgondolkodott, és elhatározta, hogy B-nek elárulja az információt, C-nek kíméletből nem szól semmit. B maga elé nézett, és így sóhajtott:

– Most már csak ketten vagyunk, akik reménykedhetünk a kegyelemben, te és én. Az esélyünk tehát 1/3-ról ½-re nőtt, mégis ugyanolyan nyugtalan vagyok, mint azelőtt!

– Megértelek – válaszolta A –, ráadásul én még abban sem vagyok biztos, hogy az esélyeim javultak volna. Az őr ugyanis énrólam nem nyilatkozott… (Hiszen tilos neki nyilatkozni, és én ravaszul úgy tettem föl a kérdésemet, hogy ne kelljen megszegnie a szabályt.) Ezért a válasza csak rád és C-re vonatkozott. Az én esélyem tehát továbbra is egyharmad. Ha valakinek nőtt az esélye, az csak te lehetsz. Szívesebben lennék most a helyedben!

– Úgy gondolod, jobban járnál, ha az én „B" feliratú rabruhámat húznád magadra? – kérdezte B, fejét ingatva. – Semmi értelme nem lenne. Teljesen mindegy, hogy az őr kinek a kérésére árulta el C kivégzését. A lényeg az, hogy C-t biztosan kivégzik, és ha kivégzik, akkor a mi esélyeink megváltoznak (szerencsére megnőnek), de egyenlőek maradnak, hiszen eredetileg is egyenlőek voltak.

Bár suttogva vitatkoztak, az őr meghallotta, és együtt érző hangon így szólt:

– Látom, szorult helyzetetekben is élvezetet okoz nektek az esélyek latolgatása és matematikai elemzése. De a helyzet az, hogy a vitátoknak semmi értelme. A sorsolás tegnap volt. Előtte mindenki esélye 1/3 volt. A sorshúzás után azonban már nincsenek esélyek, a kérdés eldőlt. Én pontosan tudom, melyikőtök kap kegyelmet. A ti vitátok ezen már nem változtathat. Nagyon sajnálom.

Kinek van igaza, A-nak, B-nek vagy az őrnek? Igaz-e, hogy A esélye nem változott? Vagy valóban ½-re nőtt? Érdemes volna-e A-nak ruhát cserélnie B-vel? Igaz-e, hogy a sorshúzás után nincs is min vitatkozni?

Alább sorra vesszük, hogy miért nem jó a fentebbi válaszok *egyike sem*. És választ keresünk arra, miért marad még sok matematikában-jártas emberben is kétely a helyes megoldás ismerete után is?

* *Főiskolai tanár, tanszékvezető, Általános Vállalkozási Főiskola*

Igaz-e, hogy A esélye ½-re nőtt?

Az ½-re vezető érvelés – amelyet a feladatban B képviselt – részletezve a következő:

Jelölje A, B és C azt az eseményt, hogy az illető rab kegyelmet kapott. A felülvonás azt jelenti, hogy az illető rab *nem* kap kegyelmet. A $\overline{C}$ esemény tehát azt jelenti, hogy a C rabot kivégzik..

Ezen érvelés szerint a három rab sorsa együttesen háromféleképpen alakulhat, mindegyik változatnak azonos az esélye. A lehetséges esetek:

$\overline{A}\ \ \overline{B}\ \ C$ Ha C a szerencsés rab, az őr B-re mutat, és A meghal.
$\overline{A}\ \ B\ \ \overline{C}$ Ha B a szerencsés rab, az őr C-re mutat, és A meghal.
$A\ \ \overline{B}\ \ \overline{C}$ Ha A a szerencsés rab, az őr akár B-re, akár C-re mutat, A megmenekül.

Ezek szerint két olyan (egyenlő esélyű) eset van, amikor C rabot kivégzik, ezek közül az egyikben A rab megmenekül. Tehát ha A számára kiderült, hogy C-t kivégzik, akkor az esélyei 50%-ra nőttek.

Formálisan egy feltételes valószínűség, a $P\left(A\middle|\overline{C}\right)$ értéke a kérdés: „Mi a valószínűsége az A rab megmenekülésének, ha kiderült, hogy a C rab biztosan nem menekül meg?” A feltételes valószínűség definíciója szerint:

$$P\left(A\middle|\overline{C}\right)=\frac{P\left(A\overline{C}\right)}{P\left(\overline{C}\right)}=\frac{\frac{1}{3}}{\frac{2}{3}}=\frac{1}{2}$$

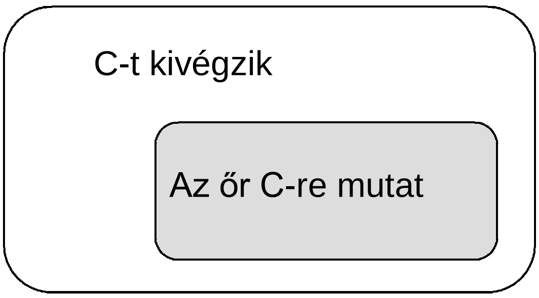
Az ½-re vezető gondolatmenet hibája

A fenti érvelés azonban hibás. Az benne a probléma, hogy *„C-t kivégzik” és „Az őr C-re mutat”* események nem azonosak egymással. Ha az őr C-re mutat, akkor C-t biztosan kivégzik, ennek esélye 100%. Fordítva azonban nem igaz: ha C-t biztosan kivégzik, abból nem következik, hogy az őr C-re fog mutatni. A két eseményhez tartozó valószínűség általában nem azonos. Márpedig ami a rabok számára információ, az az őr rámutatása.

Jelölje M_C azt, hogy az őr a C rabra mutat. Tehát: $M_C \neq \overline{C}$.

A két esemény közötti kapcsolat: $M_C \subseteq \overline{C}$, szemléletesen:

A kisebb szürke téglalap a nagyobbik fehérnek a része (részhalmaza). Ha a kis téglalapban állunk,



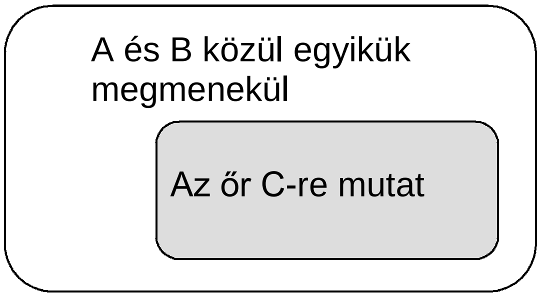
akkor benne vagyunk a nagyobbik téglalapban is, de ha csak annyit tudunk állítani, hogy benne vagyunk a nagyban, akkor abból természetesen nem következik, hogy a kicsiben is benne lennénk. Azaz: ha az őr C-re mutat, akkor C-t biztosan kivégzik, de abból, hogy C-t kivégzik, nem következik, hogy az őr C-re mutat.

Ezt a továbbiakban már nem hagyjuk figyelmen kívül, de a nehézségek sora ezzel még nem ért véget.

Igaz-e, hogy 1/3 marad A rab megmenekülési esélye?

Vezessük be, tehát az M_C eseményt, ami azt jelenti: „Az őr C-re mutat". Ekkor a feladat a $P(A|M_C)$ feltételes valószínűség meghatározása: „Mi a valószínűsége annak, hogy A megmenekül, feltéve, hogy az őr C-re mutat?"

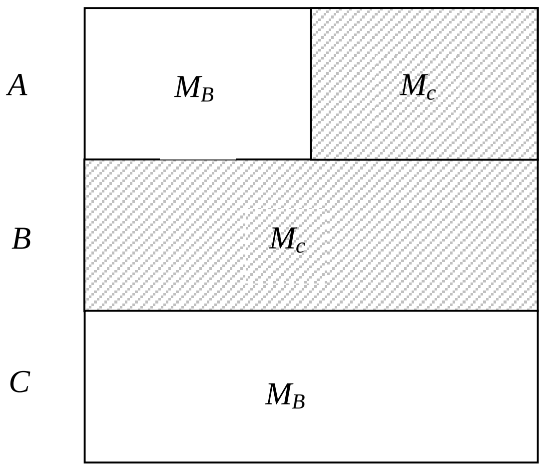
Ha C-t biztosan kivégzik, akkor A és B közül egyikük biztosan megmenekül. Az ábránk ennek megfelelően átírható.



Ábránkat két vonatkozásban is tovább fejleszthetjük.

- Egyrészt: tovább tagolhatjuk az ábrát, elkülöníthetünk az A, B és C eseményeknek egy-egy téglalapot, és mindegyiken belül kijelölhetünk egy M_C-nek megfelelő rész-téglalapot.
- Másrészt: elkészíthetjük úgy a Venn diagramunkat, hogy a részterületek az események *valószínűségével* legyenek arányosak. Az egész terület célszerűen egy egységnégyzetbe foglalható, tekintettel arra, hogy az A, B, és C események teljes eseményrendszert alkotnak.

Így az alábbi szemléletes ábrát kapjuk. Az M_C eseményt a szürke terület jelzi, ennek komplementere a M_B, ami az ábrán fehér terület. A *függőleges oldalhosszak* az A, B, C események valószínűségével, *vízszintes oldalak* a feltételes valószínűségekkel egyeznek meg. Az ábrából a feltételes valószínűség fogalma, a teljes valószínűség tétele és a Bayes tétel is könnyen megérthető.



Az ábra alakjára, méreteire vonatkozóan a következőket vettük figyelembe:

- Az A, B, C téglalapok egyformák: mindegyik alapja 1, magassága 1/3, mert P(A) = P(B) = P(C) = 1/3
- A C téglalap egésze fehér, nincs benne szürke terület, mert ha C menekül meg, akkor az őr nem mutathat C-re. Formálisan: $P(M_C|C) = 0$.
- A középső B téglalap egésze szürke, mert ha a B rab az, aki megmenekül, akkor az őr biztosan C-re fog mutatni, másra nem mutathat, tehát $P(M_C|B) = 1$.
- Hátra van még, hogy az A téglalapban mennyi a szürke rész aránya, azaz mennyi a $P(M_C|A)$ értéke. Ez nincs megadva a feladatban. Gondolhatjuk azt, hogy az őr – az A rab kegyelme esetén – véletlenszerűen választ B és C között, ez esetben $P(M_C|A) =$ ½, és ekkor az A téglalap felerészben lesz szürke, a fenti ábrát így rajzoltuk meg.

Végső soron a két eseményrendszer *együttes* valószínűség-eloszlását szemlélteti az ábra. Az ábra alapjául szolgáló adatokat az alábbi táblázat foglalja össze:

	M_B	**M_C**	
A	1/6	1/6	**1/3**
B	0	1/3	**1/3**
C	1/3	0	**1/3**
	½	½	**1**

Az ábrából a következő olvasható le: Az A téglalapon belüli szürke rész aránya az összes szürke területnek az 1/3-a. Ez azt jelenti, hogy *azon esetek közül, amikor az őr a C-re mutat, 1/3 azon esetek aránya, amikor A megmenekül.* Képlettel: $P(A|M_C) =$ 1/3. Tehát ha az őr C-re mutat, akkor A megmenekülési esélye 1/3. Az eredmény meglepő, mondhatni, paradox, hiszen ez ugyanannyi, mint a rámutatás előtt!

Amit az ábráról leolvastunk, az képletekkel így írható föl:

$$P(A|M_C) = \frac{P(M_C|A)\,P(A)}{P(M_C|A)\,P(A) + P(M_C|B)\,P(B) + P(M_C|C)\,P(C)}$$

Ez a *Bayes-tétel*, amit szemléletes tartalma alapján fogalmaztunk meg. A számlálóban az A téglalapon belüli szürke terület nagysága van, a nevezőben a szürke terület *teljes* nagysága. A példabeli adatokkal:

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

Az 1/3-os válasz problémája

Az 1/3-os válasz tehát helyesnek látszik, bár a feladatbeli érvelés, miszerint A esélye azért marad 1/3, mert az őr válasza nem A-ra vonatkozott, és ezért csak B és C viszonylatában van jelentősége, természetesen helytelen. Ebben a gondolatmenetben az a probléma, hogy a $P(M_C|A)$-nek nem kell szükségképpen ½-nek lennie. A feladat erről nem mond semmit. Természetesen lehetséges, hogy amikor az őrnek döntenie kell B és C között, akkor az őr feldob egy szabályos pénzérmét. Ez esetben $P(M_C|A)$ = ½ . De lehetségesek egyéb kézenfekvő esetek is:

- Tegyük fel például, hogy az őr A kegyelme esetén *betűrend szerint* B és C közül B-re fog mutatni, mert a listáján B előbb következik. Ekkor $P(M_C|A) = 0$. Ez azt jelenti, hogy ha az őr C-re mutat, akkor A biztosan nem kap kegyelmet.
- Vagy fordított betűrend alapján biztosan C-re fog mutatni. Ez esetben $P(M_C|A) = 1$, és a végeredmény ½, tehát A esélye nőtt. Adott feltevések mellett tehát az ½-es eredmény is lehet jó megoldás.
- Vagy az őr földob egy érmét, és fej esetén fog C-re mutatni. De az érme kopott és görbe, a fej valószínűsége 60%, tehát $P(M_C|A) = 0{,}6$, és a végeredmény 37,5%.

Ha pedig tartjuk magunkat ahhoz, hogy a C-re mutatás valószínűségét nem tudjuk, akkor csak annyit mondhatunk, hogy A esélye valahol 0 és ½ között lehet.

Érdemes lenne-e A-nak ruhát cserélni B-vel?

A fentiek alapján erre egyértelmű választ tudunk adni. A-nak *általában* érdemes ruhát cserélnie, mert A esélye legfeljebb ½, B esélye pedig legalább ½ (B esélye ugyanis A esélyének a komplementere.). A ruhacsere **A** esélyeit általában javítja, és semmiképp se rontja.

Mi okozza a feladat nehézségét?

Most már tudjuk a megoldást. Vizsgáljuk meg, mi okozta ebben a feladatban a nehézséget? A népszerű logikai feladatok többnyire olyanok, hogy első ránézésre nehéznek, vagy akár megoldhatatlannak látszanak, aztán amikor nagy nehezen rájövünk a megoldásra, akkor a megoldás egyszerűnek és kézenfekvőnek tűnik. *A három rab esete* nem ilyen. Ez első hallásra egyszerűnek tűnik, és aztán kiderül, hogy nem is olyan egyszerű. A kézenfekvőnek tűnő megoldás nem jó, illetve egyidejűleg többféle – egymásnak ellentmondó – „kézenfekvő" megközelítés is helyesnek látszik. És amikor végre tisztáztuk a megoldást, akkor sincs olyan érzésünk, hogy most már minden rendben, minden világos. Sőt, a megoldás ismeretében is hajlandóak vagyunk elbizonytalanodni, és matematikában járatos emberek is sokszor kétkedve fogadják a megoldást.

Az egyik nehézséggel már találkoztunk. Ez abban állt, hogy a C és az M_C eseményeket meg kellett különböztetni egymástól. De önmagában még ez sem elég. Hiszen akad, aki megkülönbözteti a kettőt, de azt gondolja, hogy elegendő abból a tényből kiindulni, hogy C-t kivégzik, mert hiszen ha az őr – úgymond – C-re mutatott, akkor C-t biztosan kivégzik. És ez utóbbi állítás önmagában igaz is! Mégsem helyes abból a tényből kiindulni, hogy C-t kivégzik. Mert itt nem egyszerűen C kivégzése a további következtetések alapja, hanem maga az $M_C \Rightarrow C$ összetett állítás. Végül is az $P(M_C|A)$ feltételes valószínűségből kell következtetni az $M_C \Rightarrow C$ állítás figyelembevételével egy másik feltételes valószínűségre, a $P(A|M_C)$-re.

A másik leginkább meglepő mozzanat minden bizonnyal az, hogy a feladatban elhangzik egy lényegi információ, és A esélye mégsem változik. Ezek szerint ez még sem volt lényegi információ, és a tankönyvi megoldások ezt szeretik is hangsúlyozni.

Maradjunk a $P(M_C|A)$ = ½ esetnél. Ekkor azt kaptuk, hogy

$$P(A|M_C) = P(A) = \frac{1}{3}$$

Az A esemény valószínűsége önmagában ugyanannyi, mint az A eseménynek az M_C feltétel melletti valószínűsége. Más szóval: az A és az Mc események *függetlenek*. Ami viszont többszörösen is sajátos eset, mert:

- először is, mint fentebb tisztáztuk, ha P(Mc/A) nem ½, akkor az A és az Mc események nem *függetlenek;*
- másrészt, ha az A és az Mc események *függetlenek* is, az $\{A, B, C\}$ és a $\{M_B, \ M_C\}$ teljes eseményrendszerek már *nem függetlenek* egymástól.

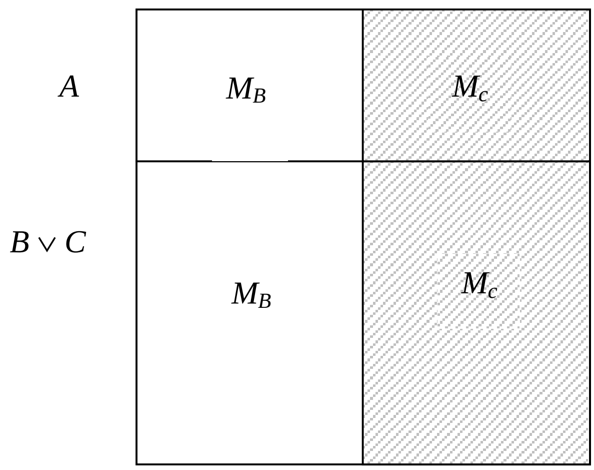
Az ábránkról ránézésre leolvasható, hogy a két eseményrendszer valóban nem független. ***Függetlenség esetén ugyanis – amint az könnyen belátható –a szürke M_C terület aránya minden egyes téglalapon belül azonos lenne***.

Önmagában és általában teljesen jogos tehát az a várakozás, hogy az őr rámutatása alapján a szereplők esélyei módosulni fognak. És ebben a példában is módosulnak, hiszen C reményei megsemmisülnek, B esélyei is javulnak, és csak egy nagyon speciális körülménynek köszönhető, ha A esélyei változatlanok maradnak, más feltételek érvényesülése esetén nem maradnak változatlanok, nőhetnek is, csökkenhetnek is, mint láttuk.

Ha $P(M_C|A) = {}^1/_2$ és összevonjuk a B és C eseményeket, tehát ha csak az A és az M_C eseményeket nézzük, azaz, az $\{A, \bar{A}\}$ és az $\{M_B, \ M_C\}$eseményrendszereket, akkor az ennek megfelelő táblázatból és ábrából jól látható hogy ez a két eseményrendszer egymástól valóban független.

	M_B	M_C	
A	1/6	1/6	**1/3**
$\bar{A}$	1/3	1/3	**2/3**
	½	½	**1**

Az ennek megfelelő ábra pedig:



Ha ez utóbbi ábrában az A téglalapon belül az M_C arányát növeljük vagy csökkentjük, (tehát $P(M_C|A)$ értékét változtatjuk, akkor természetesen a függetlenség már ebben az összevont formában sem fog fennállni.

Megjegyzendő, hogy a tankönyvekben nem elég súllyal, vagy egyáltalán nem szerepel *teljes eseményrendszerek* közötti függetlenség, aminek a speciális eseteként lenne értelmezhető két *egyedi esemény* függetlensége. Pedig a statisztikában az *asszociáció*, tehát két minőségi ismérv közötti kapcsolat szorossága, továbbá *két valószínűségi változó* függetlensége is erre a megközelítésre épül.

Az őr véleménye

Az őr azt állítja: miután a sorsolás lezajlott, nincs már tovább értelme valószínűségről beszélni, hiszen 100%, hogy a kisorsolt rab kapja a kegyelmet. Az őr állítására információelméleti megközelítéssel lehet jól válaszolni. Mert igaz ugyan, hogy a kiválasztás eldőlt, de a rabok *bizonytalansága* megmaradt. És a rabok bizonytalansága szempontjából egyre megy, hogy

- sorsolás előtt vannak;
- vagy utána vannak, de még nem tudják az eredményt.

És akkor érkezik egy információ. Ha az érkezne, hogy kinek kegyelmeznek meg, akkor a bizonytalanság teljesen megszűnne. Ennek az információnak az értéke 1,6 bit lenne, mert az eredeti helyzet bizonytalansága, azaz entrópiája kereken 1,6. Az R a kegyelem valószínűségének *eloszlása* a rabok között. H jelöli az entrópiát.

$$H(R) = -p(A)\log_2 p(A) - p(B)\log_2 p(B) - p(C)\log_2 p(C)$$

$$H(R) = -\frac{1}{3}\log_2\frac{1}{3} - \frac{1}{3}\log_2\frac{1}{3} - \frac{1}{3}\log_2\frac{1}{3} = 1{,}585 \text{ bit}$$

Az őr azonban nem szünteti meg teljes egészében a bizonytalanságot. Az ő információja kevesebb, mint 1,6 bit, csak 2/3 bit (0,67 bit), tehát az őr információja utáni helyzet entrópiája kereken csak 0,9 bit.

$$H(R|M_C) = -\frac{1}{3}\log_2\frac{1}{3} - \frac{2}{3}\log_2\frac{2}{3} = 0{,}918 \text{ bit}$$

Az információnyereség, illetve a rámutatás információmennyisége:

$$H(R|M_C) - H(R) = 1{,}585 - 0{,}918 = \frac{2}{3}$$

A rámutatás tehát a rabok számára 2/3 bit információt nyújtott. Nem az *egyes* rabok számára, hiszen az entrópia nem egyes eseményekre, hanem az eloszlásra vonatkozik.

Az *egyes* rabok helyzete eltérően alakul: C számára például egyértelművé vált a helyzet, az ő számára az 1/3 esély a 0-ra csökkent. Az őr rámutatása C-nek 0,918 bitet ért, amennyiben a {C, nem C} teljes eseményrendszer entrópiáját vesszük alapul.

Mint láttuk (ha a C-re, illetve a B-re mutatás esélye egyenlő) akkor A számára a helyzet nem változik. Számára az őr rámutatása 0 információt jelent.

B számára a bizonytalanság, ha a (B, nem B) eseményrendszert vesszük, szintén nem változik, mert az (1/3, 2/3) eloszlás entrópiája ugyanaz, mint a (2/3, 1/3) eloszlásé. Ez maga is paradox következtetés: B *esélye* 1/3-ról 2/3-ra nőtt, miközben helyzetének *bizonytalansága* nem változott.

Ez azt mutatja, hogy az információmennyiség általános esetben fontos mutató a helyzet értékelésére, de például B helyzetének változása a kegyelem *valószínűségének változásával* jobban jellemezhető, mint az őr által nyújtott információ mennyiségével.

FELHASZNÁLT ÉS AJÁNLOTT IRODALOM

Ferenczy Miklós: Valószínűség-számítás és alkalmazása. Feladatgyűjtemény. Nemzeti Tankönyvkiadó. Budapest, 1998.

Raymond Smullyan: Seherezádé rejtélye – és más bámulatos fejtörők, régiek és újak. Typotex, Budapest, 1999.

Rényi Alfréd: Napló az információelméletről. In: Rényi Alfréd Ars Mathematica. Typotex Kiadó, Budapest, 2005.

Székely J. Gábor: Paradoxonok a véletlen matematikájában. Typotex, Budapest, 2004.

Internetes címek a 3 rab esetéről és a rokon Monty Hall paradoxonról:
http://home.att.net/~numericana/answer/counting.htm#3pp
http://www.comedia.com/hot/monty.html
http://phatnav.com/wiki/index.php?title=Monty_Hall_problem