Információs rendszerek elméleti alapjai

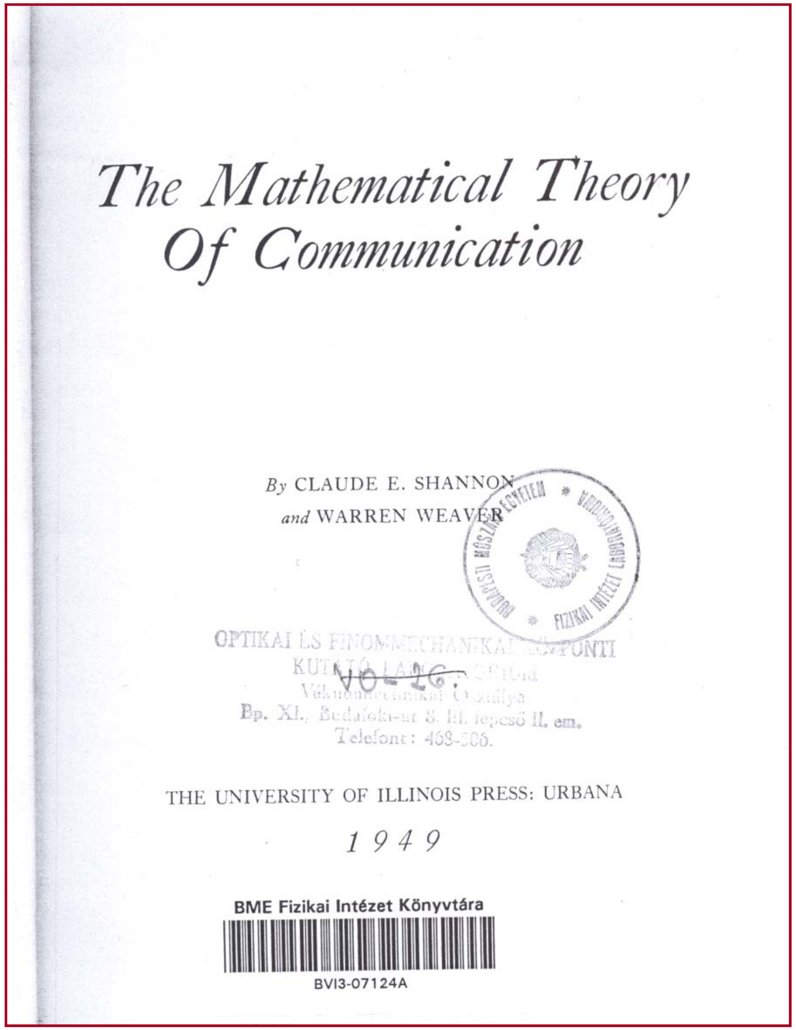
Információelmélet

Irodalom

Irodalomjegyzék

1. J. Aczél, Z. Daróczy. *On Measures of Information and Their Characterization*. Academic Press, New York. 1975.
2. R. B. Ash. *Information Theory*. Interscience, New York. 1965.
3. J. Berstel, D. Perrin. *Theory of Codes*. Academic Press, New York. 2002.
4. G. Birkhoff, T.C. Bartee. *A modern algebra a számítógéptudományban*. Műszaki Könyvkiadó, Budapest. 1964.
5. T. M. Cover, J.A. Thomas. *Elements of information theory*. Wiley, New York. 1991.
6. Csiszár I., Fritz József. *Információelmélet*. Tankönyvkiadó, Budapest. 1980.
7. C. E. Shannon, W.Weaver. *A kommunikáció matematikai elmélete*. OMIKK, Budapest. 1986.
8. Györfi László, Győri Sándor, Vajda István: Információelmélet és kódelmélet, Typotex, 2000
9. Linder Tamás, Lugosi Gábor: Bevezetés az információelméletbe, Tankönyvkiadó 1990

Az információelmélet kezdetei



The Mathematical Theory Of Communication

By CLAUDE E. SHANNON
and WARREN WEAVER

THE UNIVERSITY OF ILLINOIS PRESS: URBANA

1 9 4 9

Információlemélet alapjai

1. Kommunikáció
   1. Információs rendszerek világa
   2. Véletlen jelenségekre épít
2. Shannon féle entrópia
   1. Jövőre vonatkozik
3. Kolmogorov entrópia
   1. Számításelméletre épül, múltra vonatkozik.

Kommunikáció lényege emberi elmék (tudatok) kölcsönhatása.

Shannon hírközlési modellje

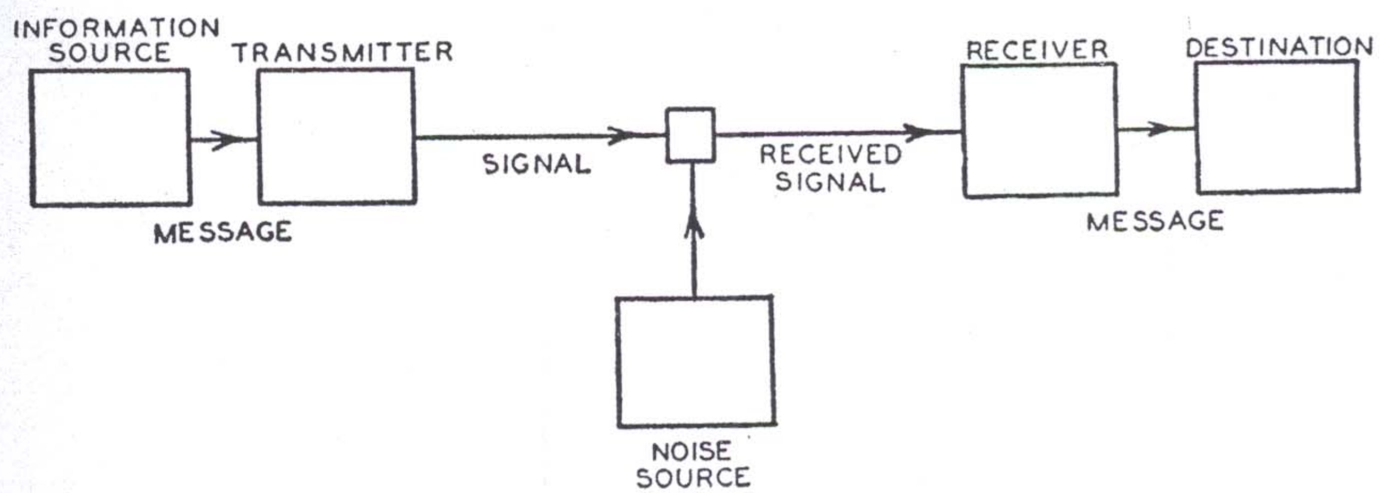
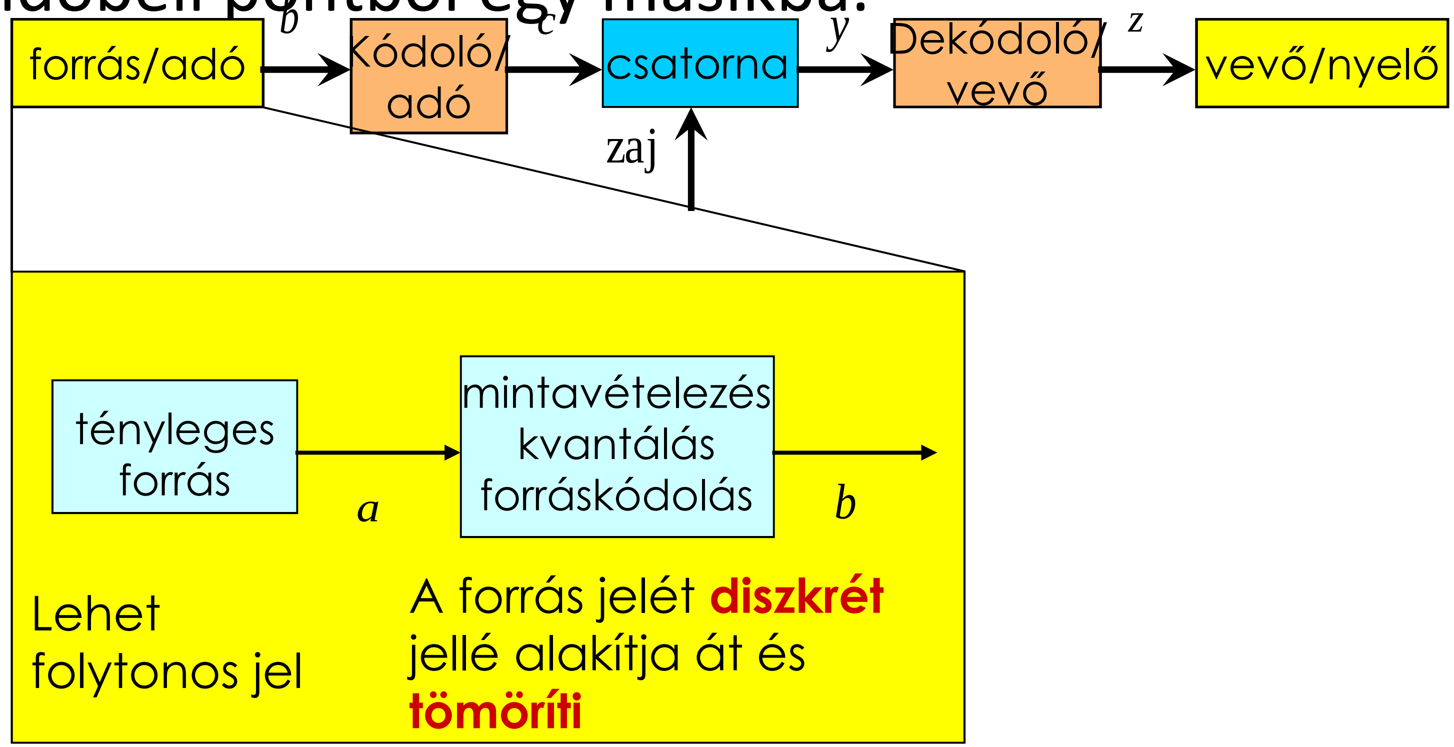


Fig. 1—Schematic diagram of a general communication system.

információforrás	adó	csatorna	vevő	rendeltetési hely
		zajforrás		Címzett

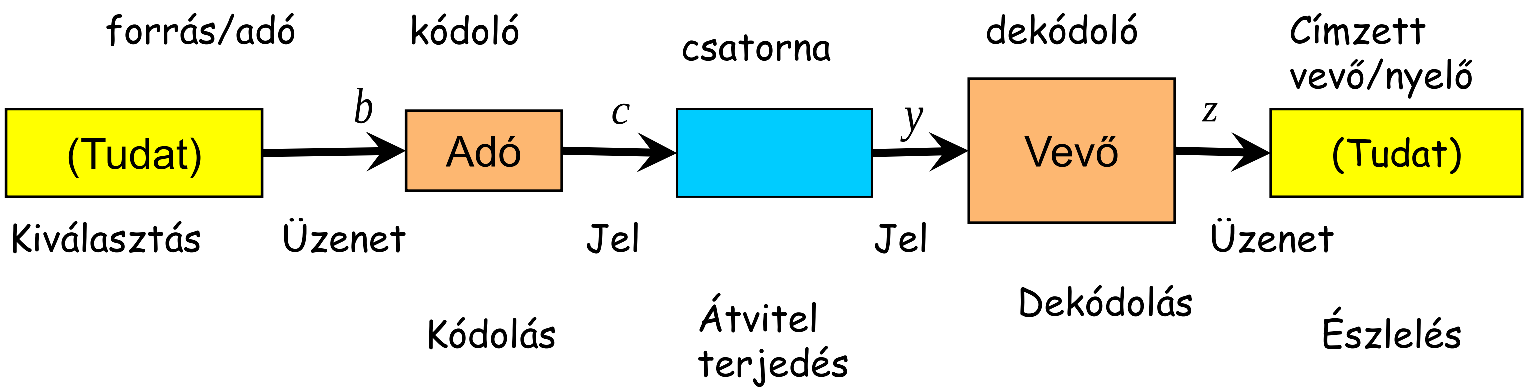
Shannon hírközlési modellje

A hírközlés során egy üzenetet juttatunk el egy tér- és időbeli pontból egy másikba.



Shanon kommunikációs modellje

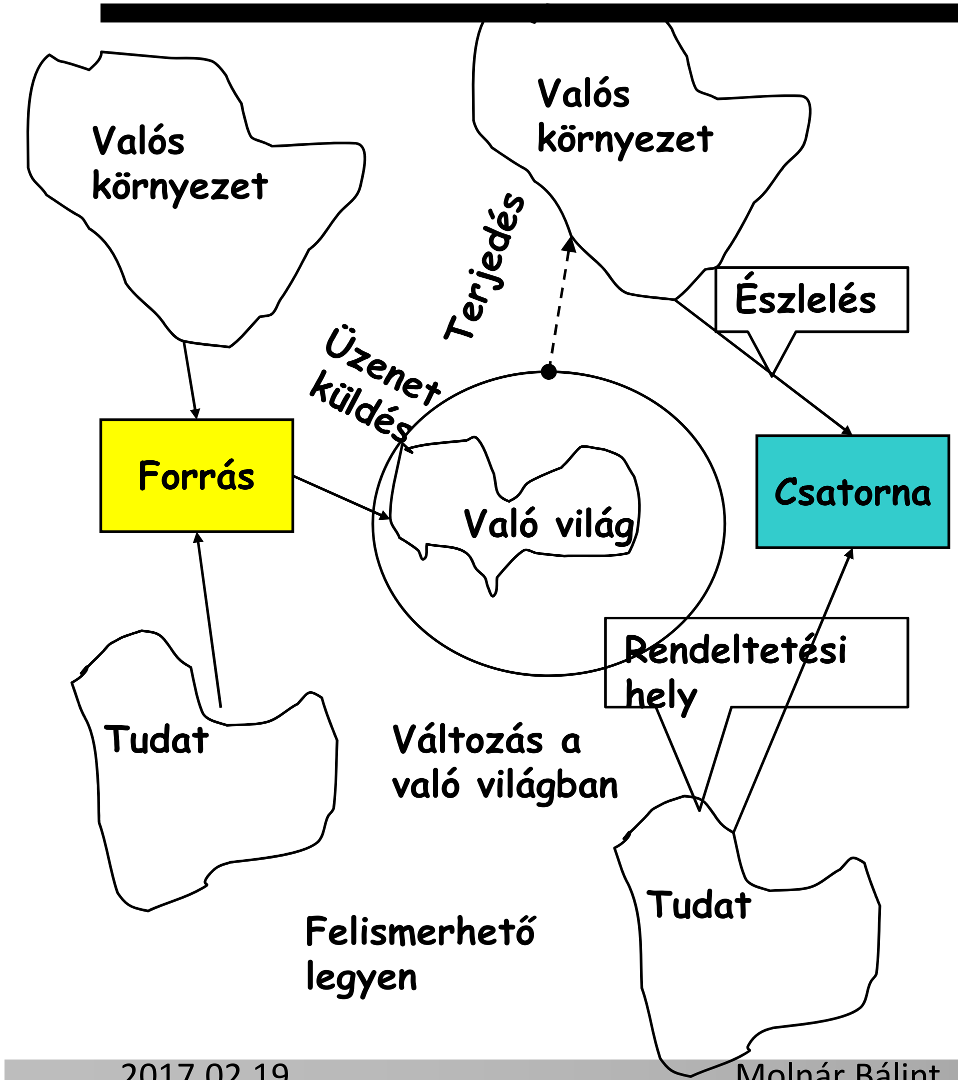
W. Weaver : A kommunikáció felöleli mindazokat az eljárásokat, amelyekkel az egyik emberi elme a másikra hatni képes



Szintek

1. ***Mennyiségi szint***:
   1. Mennyire vihető át rekonstruálhatóan a kibocsátott üzenet az adott csatornán (mennyiségi jellemzők, és csatorna kapacitás).
   2. ***Megértési szint*** (jelentés, szemantika).
   3. ***Hatékonysági szint*** (a kívánt hatás elérése, *„hatásosság, eredményesség; effectivity/efficacy*”; szemiotika, pragmatika).

Kommunikáció modellje



Valamit észlelünk a valós környezetben; kapcsolatba kerül a *tudatunkkal*; üzenet kibocsátás történik;változást vált ki a *valós világban*; észleli a címzett; kapcsolatba kerül a *tudattal*.

A *terjedés* időben történik -> torzulások lehetnek.

A kommunikáció működőképességének alapvető feltétele, hogy különböző üzenetekhez olyan különböző változás legyen létrehozható a valós világban, hogy az észlelő is különbözőnek ismerhesse fel.

Hatékonyság: emlékek megőrződése.

Történeti állomások

- Beszéd kialakulása
  - Hangképzés, megfelelő finomságú, hallás, tudat/agy, gesztikulálás, testbeszéd
  - Hanghullám terjedése
  - A közös ismeretek egy része van mindenkinél, csak élő tudatban létezik
  - Egy részét elérhetővé tesszük mások számára.
- Feljegyzés – írás, rajz, kép
  - Írás megmarad, tartósság [jövő számára tudattól függetlenül megőrződik]
  - Nyomtatás (könnyű sokszorosítás)
    - Középkor, kódexek
    - Irattár, levéltár
    - Könyvtár, szolgáltatás
    - Rögzített üzenetek azonos formában
    - Többszörösen, elérhetővé válik
    - Közös ismeretetek,
    - Bürokrácia

Az információelmélet kezdetei

- Állatvilágról
- Emberi információcsere fontosabb lépései
  - közvetlen információcsere
  - távoli személyek közti információcsere
  - Gépek közvetítésével való hírközlés
  - gépek által generált adatforgalom
- Az információelmélet kezdetei, Hartley, Shannon
- Az információelmélet a gyakorlatban

Jelek

Rádiófrekvenciás műsorszórás -> hírközlés

Telefon ->adatátvitel, távközlés (telematika)

Fénykép-> digitális műszerek

Írógép

Számítógép -> adattárolás, adatelérés, számítások (érzékszervek jelentősége csökken)

Közös észlelés: nem szükséges, hogy egy időben történjen mindenkinél.

Zajos csatornák

modem → phone line → modem

Galileo → radio waves → Earth

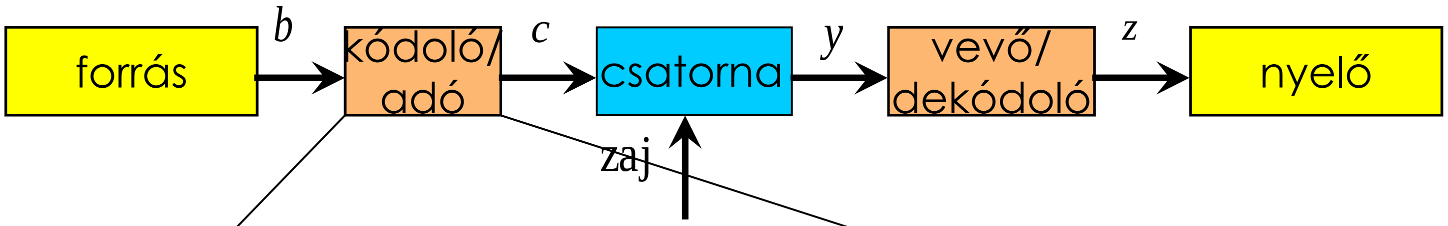
parent cell → daughter cell

parent cell → daughter cell

computer memory → disk drive → computer memory

Shannon hírközlési modellje

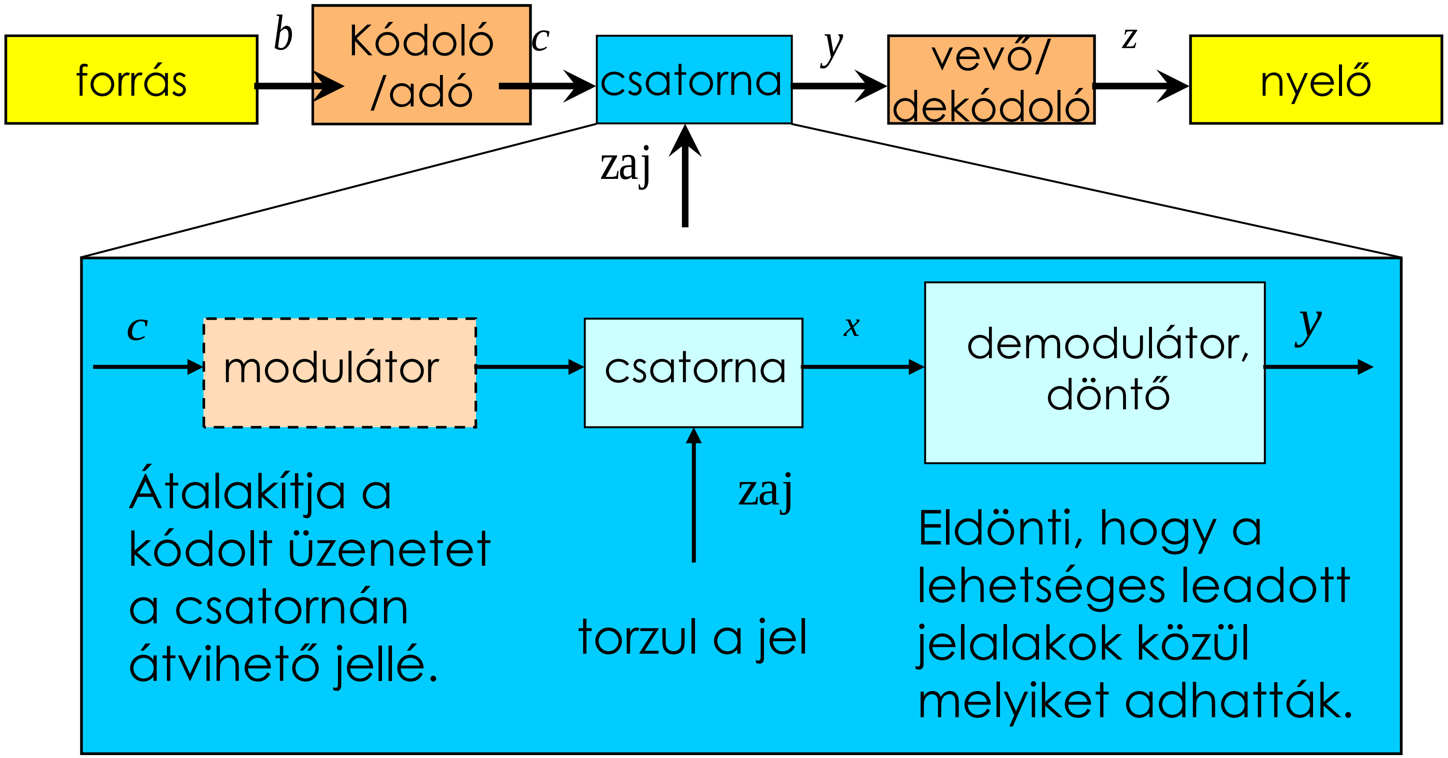
A hírközlés során egy üzenetet juttatunk el egy tér- és időbeli pontból egy másikba.



Csatornakódolás avagy hibajavító kódolás:
lehetővé teszi a zajos csatornán való biztonságos(abb) üzenetátvitelt, a keletkező hibák jelzését kijavítását.

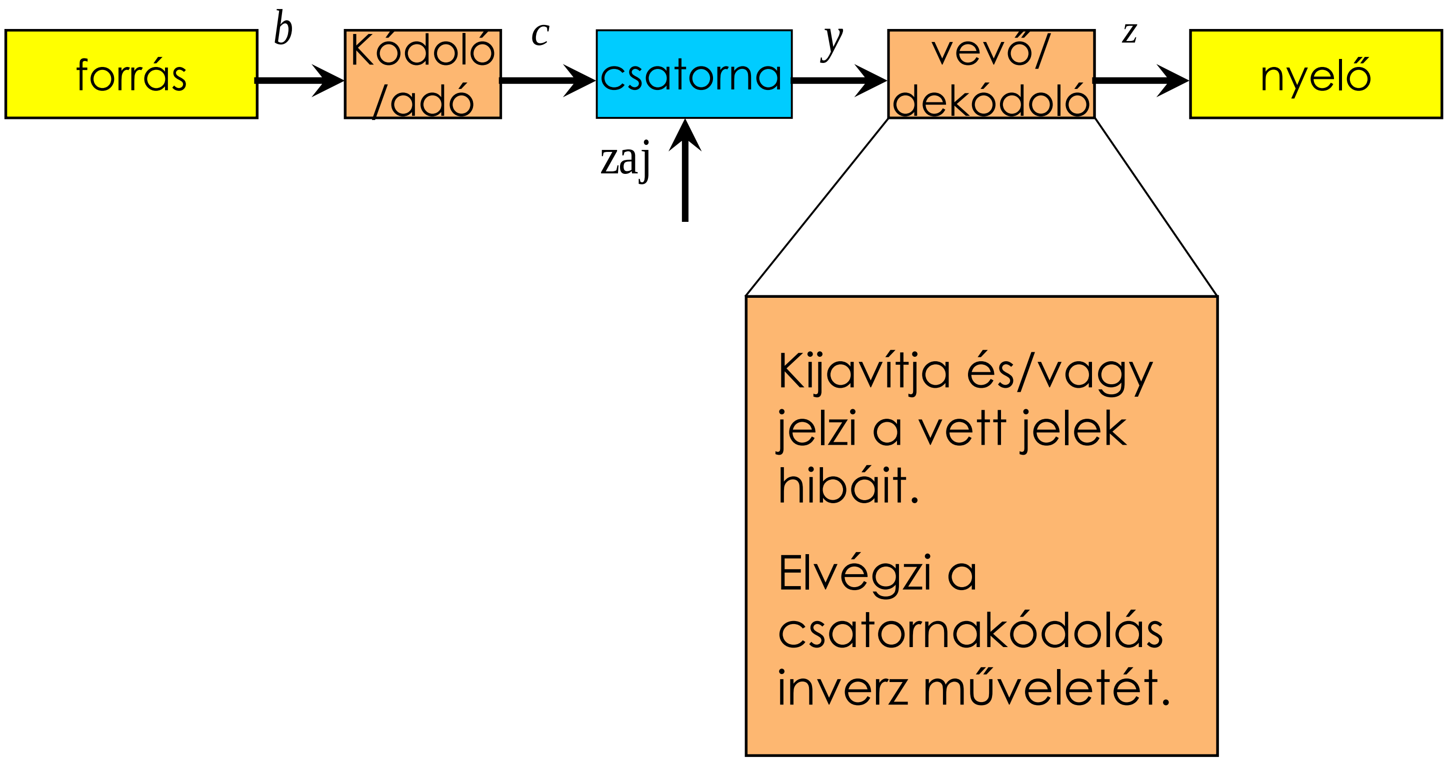
Shannon hírközlési modellje

A hírközlés során egy üzenetet juttatunk el egy tér- és időbeli pontból egy másikba.



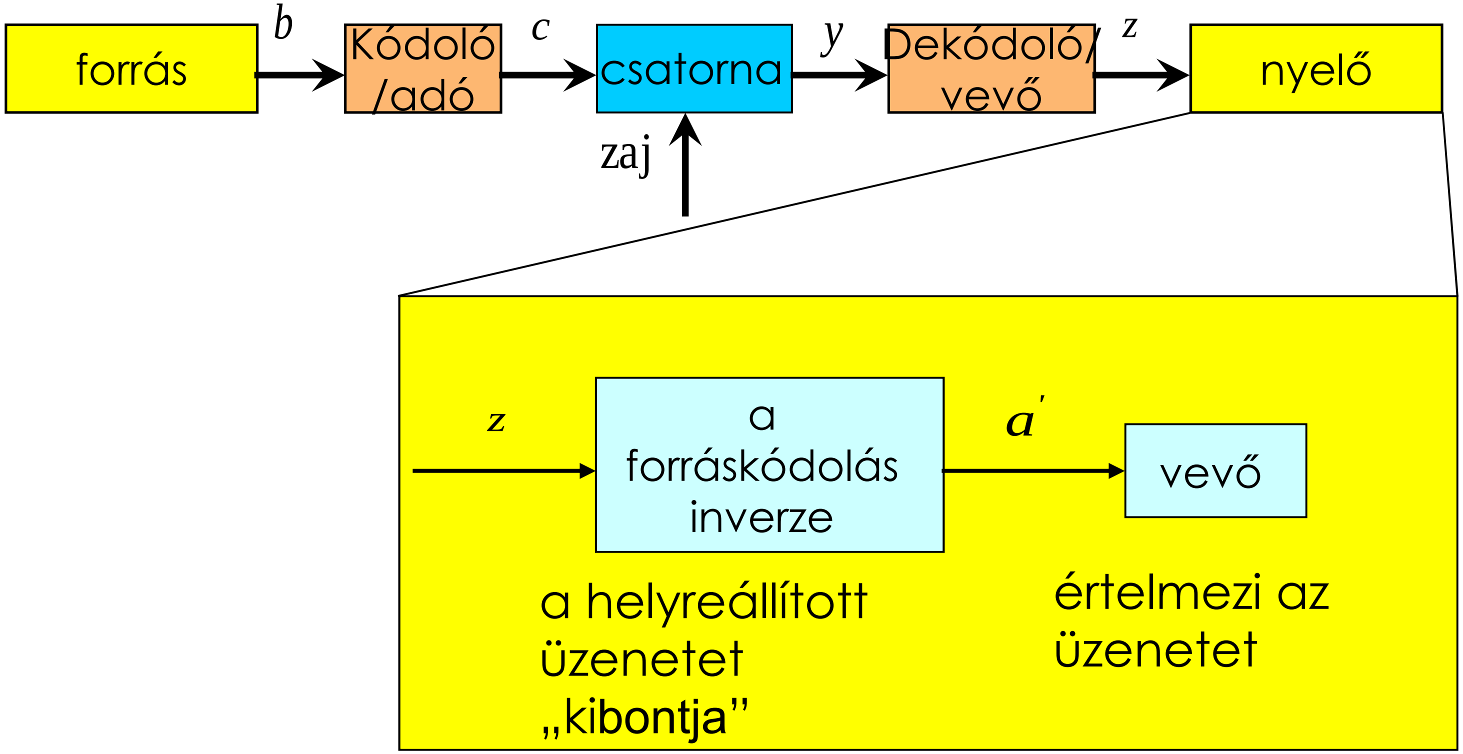
Shannon hírközlési modellje

A hírközlés során egy üzenetet juttatunk el egy tér- és időbeli pontból egy másikba.



Shannon hírközlési modellje

A hírközlés során egy üzenetet juttatunk el egy tér- és időbeli pontból egy másikba.



a vevő a csatorna végén levő "műszaki" berendezés, leemeli a jeleket a csatornáról, ezeket fogja a dekódoló az üzenetnek megfelelő szimbólumokká alakítani

Matematikai kitérő – A valószínűségszámítás alapfogalmairól

Esemény: sokszor végrehajtható kísérlet eredménye.

Kísérlet, ill. kimenetel lehet például:

- Egy dobozból golyókat veszünk ki, és vizsgáljuk a színüket.
  Esemény, hogy piros, zöld, lila, … golyót találtunk.
- Fej vagy írást játszunk egy érmével.
  Esemény: a fej vagy az írás oldal van felül.
- Minden kedden reggel 7 és 8 óra között vizsgáljuk egy buszmegállóban a megálló buszok számát.
  Esemény: 0 busz állt meg, 1, 2, … busz állt meg.
- Egy hírközlési csatornára bocsátunk egy bizonyos jelet és vizsgáljuk a kimeneti jelet.
  Esemény: a vett jel azonos a leadottal, a vett jel amplitúdója azonos a leadottéval, de a frekvenciája kétszeres, …

Matematikai kitérő –
A valószínűségszámítás alapfogalmairól

Az *A* esemény ellentett eseménye **(komplementer eseménynek)** a kísérlet minden *A*-n kívüli kimenetele. Jelölés: $\bar{A}$.

Egy *A* esemény valószínűsége: nagyon sokszor elvégezve a kísérletet (Ez a nagy számok törvénye, ezzel közelítjük az esemény valószínűségét)

$$p(A) \cong \frac{\text{azon kísérletek száma, amikor } A \text{ bekövetkezik}}{\text{az összes kísérlet száma}}$$

$$= \lim_{kisérletekszáma \to \infty} \frac{\text{azon kísérletek száma, amikor } A \text{ bekövetkezik}}{\text{az összes kísérlet száma}}$$

Matematikai kitérő –
A valószínűségszámítás alapfogalmairól

<u>A valószínűség jellemzői</u>:

- $1 \geq p(A) \geq 0$, az 1 valószínűségű esemény biztosan bekövetkezik, a 0 valószínűségű sohasem következik be.
- $p(A)+\mathrm{p}(\bar{A}) = 1$.

Matematikai kitérő – A valószínűségszámítás alapfogalmairól

A Kolmogorov-féle valószínűségi axiómákról:

Eseménytér: Az elemi események összessége: Ω; a teljes esemény

Események halmaza: S

Lehetetlen esemény: $\emptyset$.

Az S halmaz akkor és csak akkor **σ-algebra**, ha

- $\Omega \in S$ és $\emptyset \in S$,
- ha $A_i \in S$ $\forall i$-re, akkor $\sum_i A_i \in S$
- ha $A_1 \in S$ és $A_2 \in S$, akkor $A_1 \cap A_2 \in S$
- A megszámlálható metszet (szorzás) műveletre zárt (és az unió (összeadás) műveletére is) , azaz $i \in \mathbb{N}$

Matematikai kitérő – A valószínűségszámítás alapfogalmairól

A Kolmogorov-féle valószínűségi axiómákról:

Minden $A \in S$ eseményhez rendelhető egy $p(A)$ szám, az A <u>valószínűsége</u>, melyre a következők igazak:

- $0 \leq p(A) \leq 1$
- $p(\Omega)=1$
- ha $A_i \cap A_j = 0 \ \forall\ i \neq j$-re, akkor $p\left(\sum_i A_i\right) = \sum_i p(A_i)$
- *(A* és egy *B* esemény **<u>szorzatán/metszetén</u>** azt értjük, ha *A* és *B* is bekövetkezik)

Matematikai kitérő –

A valószínűségszámítás alapfogalmairól

Egy dobozban 15 sárga, 6 lila, 42 fehér és 11 kék golyó van. Mi a valószínűsége, hogy ha egyetlen golyót veszünk ki a dobozból (nem odanézve), akkor az

- kék lesz: $p(\text{kék}) = \frac{11}{15+6+42+11} = \frac{11}{74}$
- sárga lesz: $p(\text{sárga}) = \frac{15}{15+6+42+11} = \frac{15}{74}$
- nem sárga lesz: $p(\overline{\text{sárga}}) = \frac{6+42+11}{15+6+42+11} = \frac{59}{74} = 1 - p(\text{sárga})$
- kék vagy lila lesz: $p(\text{kék vagy lila}) = \frac{6+42}{15+6+42+11} = \frac{48}{74}$
- piros lesz: $p(\text{piros}) = 0$

Matematikai kitérő – A valószínűségszámítás alapfogalmairól

Egy *A* és egy *B* esemény **szorzatán** azt értjük, ha *A* és *B* is bekövetkezik.

A és *B* együttes bekövetkezési valószínűsége: $p(AB)/p(A \cap B)$.

Egy *A* és egy *B* esemény **összege/uniója** az, ha vagy *A*, vagy *B* (vagy mindkettő) bekövetkezik. Valószínűsége: $p(A+B)/p(A \cup B)$.

Az *A* és *B* események függetlenek, ha semmiféle befolyással nincs *A*-nak a bekövetkezése a *B* bekövetkezésére. Ekkor $p(AB) = p(A) \cdot p(B)$.

Matematikai kitérő – A valószínűségszámítás alapfogalmairól

Egyéb esetekben $p(AB) \neq p(A) \cdot p(B)$, csak azt lehet tudni, hogy

$$p(A+B) = p(A) + p(B) - p(AB),$$

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

és

$$p(A \cdot B) \leq \min(p(A), p(B))$$

Matematikai kitérő – A valószínűségszámítás alapfogalmairól

Egy dobozban 15 sárga, 6 lila, 42 fehér és 11 kék golyó van. Mi a valószínűsége, hogy ha két golyót veszünk ki a dobozból (nem odanézve), akkor

- az 1. kék lesz, a 2. fehér: $$\mathrm{p}\left(\left(1.\text{kék}\right)\cdot\left(2.\text{fehér}\right)\right)=\frac{11}{74}\cdot\frac{42}{73}$$
- mindkettő sárga lesz: $$p\left(\left(1.\text{sárga}\right)\cdot\left(2.\text{sárga}\right)\right)=\frac{15}{74}\cdot\frac{14}{73}$$
- valamelyik sárga lesz: $$p\left(\left(1.\text{sárga}\right)+\left(2.\text{sárga}\right)\right)=\frac{15}{74}\cdot\frac{14}{73}+\frac{15}{74}\cdot\left(1-\frac{14}{73}\right)+\left(1-\frac{15}{74}\right)\cdot\frac{15}{73}$$
- egyik sem lesz sárga: $$p\left(\left(\overline{1.\text{sárga}}\right)\cdot\left(\overline{2.\text{sárga}}\right)\right)=\left(1-\frac{15}{74}\right)\left(1-\frac{15}{73}\right)$$ $$\left(=1-p\left(\left(1.\text{sárga}\right)+\left(2.\text{sárga}\right)\right)\right)$$

Matematikai kitérő –

A valószínűségszámítás alapfogalmairól

Az is érdekes, hogy ha $\xi=A$ jelet adok, milyen $\eta=B$ kerül a csatorna kimenetére, ez $\eta=B$ -nek $\xi=A$ feltétel melletti feltételes valószínűségével, $p(\eta=B \mid \xi=A)$-val írható le:

$$p(\eta=B \mid \xi=A)=\frac{p(AB)}{p(A)}$$

Az együttes feltételes valószínűségek között fennáll:

$p(AB)=p(B)\cdot p(A \mid B)=p(A)\cdot p(B \mid A)$

Matematikai kitérő – A valószínűségszámítás alapfogalmairól

Egy dobozban 15 sárga, 6 lila, 42 fehér és 11 kék golyó van. Mi a feltételes valószínűsége annak, hogy ha az első kihúzott golyó kék volt, akkor a második

- fehér: $$p((2.\text{fehér})\,|\,(1.\text{kék})) = \frac{p((1.\text{kék})\cdot(2.\text{fehér}))}{p(1.\text{kék})} = \frac{\frac{11}{74}\cdot\frac{42}{73}}{\frac{11}{74}}$$
- kék: $$p((2.\text{kék})\,|\,(1.\text{kék})) = \frac{p((1.\text{kék})\cdot(2.\text{kék}))}{p(1.\text{kék})} = \frac{\frac{11}{74}\cdot\frac{10}{73}}{\frac{11}{74}}$$
- nem kék: $$p((\overline{2.\text{kék}})\,|\,(1.\text{kék})) = \frac{p((1.\text{kék})\cdot(\overline{2.\text{kék}}))}{p(1.\text{kék})} = \frac{\frac{11}{74}\cdot\left(1-\frac{10}{73}\right)}{\frac{11}{74}}$$

Matematikai kitérő –
A valószínűségszámítás alapfogalmairól

Ha az eseményekhez (kísérlet kimeneteihez) számértékek rendelhetők, (pl. árammérés), akkor kíváncsiak lehetünk a kísérlet eredményének **várható érték**ére. Legyen $X=\{x_1, x_2, \dots x_n\}$ számhalmaz a kísérlet kimenetének értékkészlete;

A kísérlet eredményei egy ξ valószínűségi változó értékei

A valószínűségi változók lehetnek **diszkrétek, folytonosak** vagy kevertek. Egy valószínűségi változót diszkrétnek nevezünk, ha értékkészlete véges (vagy megszámlálhatóan végtelen).

az x_1 kimenet valószínűsége $P(\xi=x_1)$, … az x_n-é $P(\xi=x_n)$. Ekkor ξ várható értéke (*súlyozott átlag*)

$$E(\xi)=\sum_{i=1}^{n} P(\xi=x_i)\,x_i=\sum_{i=1}^{n} p_i\, x_i$$

Matematikai kitérő – A valószínűségszámítás alapfogalmairól Eloszlás

$P(\xi = x_i) = p_i$, $i = 1, \ldots, n$, valószínűségeket a valószínűségi változó **eloszlásának** nevezzük.

Matematikai kitérő – A valószínűségszámítás alapfogalmairól

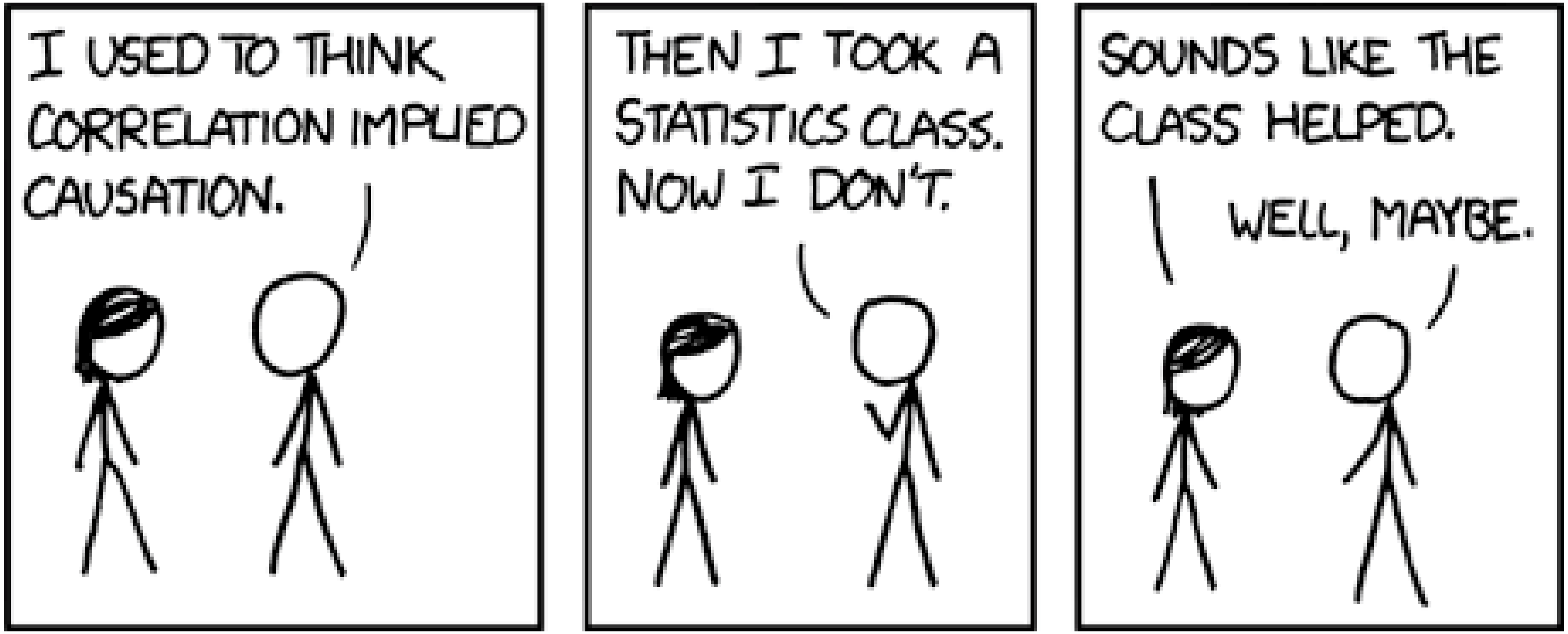
Az is érdekelhet minket, hogy átlagosan mennyire fog eltérni az eredmény a várhatóértéktől, ezt a **szórás**sal jellemezhetjük:

$$D(A) = \sqrt{E\left[\left(\xi - E(\xi)\right)^2\right]} = \sqrt{\mathbb{E}\left(\xi^2\right) - E(\xi)^2}$$

Ha több kísérletet vizsgálunk, A és B **korreláció**ja írja le az, hogy mennyire függ a kettő egymástól

$$R(\xi,\eta) = \frac{\mathbb{E}\left[\left(\xi - \mathbb{E}(\xi)\right)\cdot\left(\eta - \mathbb{E}(\eta)\right)\right]}{D(\xi)\cdot D(\eta)}$$

http://xkcd.com/552/



Az információ

Az **információ** valamely véges számú, előre ismert esemény közül annak megnevezése, hogy melyik következett be. Alternatív megfogalmazás: az információ mértéke azonos azzal a bizonytalansággal, amelyet megszüntet.

Hartley: *m* számú, azonos valószínűségű esemény közül egy megnevezésével nyert információ:

$$I = \log_2 m$$

($\log_2 m$ kérdéssel azonosítható egy elem)

Megjegyzés: Ha két eseményt egybevonunk, annak megnevezése 1 bittel lesz rövidebb, hiszen m/2 számú párból egy.

Az információ

Az **információ** valamely véges számú, előre ismert esemény közül annak megnevezése, hogy melyik következett be.
Alternatív megfogalmazás: az információ mértéke azonos azzal a bizonytalansággal, amelyet megszűntet.

Shannon: minél váratlanabb egy esemény, bekövetkezése annál több információt jelent, annál több bizonytalanságot kell kiküszöbölni.
Legyen $A=\{A_1, A_2, \dots A_m\}$ esemény-halmaz, az A_1 esemény valószínűsége p_1, … az A_m-é p_m . Ekkor az A_i megnevezésekor nyert információ:

$$I\left(A_i\right)=\log_2\frac{1}{p_i}=-\log_2 p_i \ .$$

Megjegyzés: ha $p_i=1/m$, minden i-re, visszakapjuk Hartley definícióját.

Az információ tulajdonságai

1. Csak az esemény valószínűségének függvénye.
2. Nem negatív: $I \geq 0$
3. Additív: ha $m = m_1 \cdot m_2$,
   $I(m_1 \cdot m_2) = I(m_1) + I(m_2)$

Megjegyzés: Ezek már a függvényegyenlet megoldásaként a logaritmus függvényt adják.

1. Monoton: ha $p_i \geq p_j$, akkor $I(A_i) \leq I(A_j)$
2. Normálás: legyen $I(A)$=1, ha $p(A)$=0,5. Ekkor kettes alapú logaritmus használandó és az információegysége a **bit**.
   $I(m_i)=\boldsymbol{log}\ (1/p_i)$ [bit] – binary unit/digit

Megjegyzés: ha tízes alapú logaritmust (lg-t) használunk, a hartley, az egység. Ekkor a normálás: $I(p=0,1)=1$. Ha természetes alapú logaritmussal definiáljuk az információt ($I=-\ln p$), akkor a natban mérjük az információt, a normálás pedig $I(p=1/e)=1$.

1 bit = ln 2 *nat*

1 hartley = 3,3219 bit = 2,3026 nat

1 bit = 0,6931 nat = 0,3010 hartley

1 nat = 1,4427 bit = 0,4343 hartley

Az információ

A forrásunk a következő üzenetet adta le:

„**ccndnddnnncdnndncdncdncdncdnnnncdcdncdcnnncd cccdcddcdccccnnn**”

(21 db „**c**”, 22 db „**n**”, 17 db „**d**”)

Mekkora lesz az információtartalma , ha a következő szimbólum „**c**” lesz?

$p(\mathbf{c}) = 21/(21+22+17) = 21/60 = 0{,}35$

$I(\mathbf{c}) \;=\; -\log_2 0{,}35 \;=\; -\ln 0{,}35/\ln 2 \;=\; 1{,}51$

Az információ

A forrásunk a ♈, ♉, ♊, ♋, ♌ szimbólumokat bocsátja ki $p_{♈}$=0,12, $p_{♉}$=0,37, $p_{♊}$=0,06, $p_{♋}$=0,21, $p_{♌}$=0,24 valószínűséggel. Mekkora lesz az információtartalma , ha a következő szimbólum a ♉ jel?

$$I(♉) = -\log_2 0{,}37 = -\ln 0{,}37/\ln 2 = 1{,}43$$

DISZKRÉT FORRÁS, DISZKRÉT, ZAJMENTES CSATORNA

Példa legegyszerűbb csatornával kezdjük (logikai, formális lehet)

Bináris, szimmetrikus csatorna

Jelei: 0,1

Átviteli idő: mindkettőre: 1/C *sec* (Hasonlóak a digitális csatornák).

Egy forrás: igen – nem üzeneteket választ

a) Szabályos pénzérme feldobása *c*-szer. Időegység alatt.
   a) 2^c egyforma lehetőség. Kódolás:
      a) Igen ->0
      b) Nem ->1
   b) 2^c különböző kódszó, egységnyi idő hosszú.

Bináris, szimmetrikus csatorna

2^c különböző kódszó, egységnyi idő hossz alatt.

Pl. $c=1024=2^{10}$, 2^{1024},

Bináris, szimmetrikus csatorna

Lottó húzás

- Igen: létezik ötös találat a szelvényemen
- Nem: nincs ötös találat a szelvényemen

$1024=2^{10}$, két hosszú húzássorozat eredményét milyen rövid idő alatt tudjuk csatornánk segítségével a címzetthez juttatni.

Bináris, szimmetrikus csatorna

3 eset :

Nincs ötös egyszer sem: kódszó 1, p_0=(1-*P*(ötös találat))^{1024};

Egy ötös van, *k*-dik: *10* (a kódszó) *sorszám* (10 bit, 1024=2^{10}) [log$_2$*k*]; p_1=(*1-P(ötös találat*))^{1023} $\cdot$*P*(*ötös találat*); 1023=1024-1= 2^{10}-1

Ha több ötös van : *11* (a kódszó) 1024 hosszú kódban jelöljük az ötös találatokat. (*k* bit)

A teljes sorozat p_2=1- p_0 - p_1

P(ötös találat)	1-P(ötös találat)	1-P(ötös találat)	1-P(ötös találat)	1-P(ötös találat)	1-P(ötös találat)	1-P(ötös találat)	1-P(ötös találat)	1-P(ötös találat)	1-P(ötös találat)

Bináris, szimmetrikus csatorna

Ötös lottó:

1 szelvény kitöltésével

az ötös valószínűsége = 1:43 949 268 = 0.00000275% , $\binom{90}{5}$

a négyes valószínűsége = 1:103 410 = 0.000967%

a hármas valószínűsége = 1:1 231 = 0.081%

a kettes valószínűsége = 1:44 = 2.273%

egy vagy nulla találat valószínűsége = 97.65%

Bináris, szimmetrikus csatorna

Átviteli bit mennyiség várható értéke:

$(p_0(1)+p_1(2+10)+p_2(2+1024)) \leq 2;$

Poisson eloszlás közelítése, binomiális eloszlás határeloszlása $np=\lambda$;

$\lambda = 1024 * \binom{90}{5}^{-1} = 2{,}33 * 10^{-5}$

$p_0=0{,}999977$	$p_0= e^{-\lambda}$
$p_1=2{,}33*10^{-5}$	$p_1= \lambda e^{-\lambda}$
p_2 számítási hibánál kisebb	

Matematikai kitérő –

A valószínűségszámítás alapfogalmairól

Poisson eloszlás

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

. Kifejezi az adott idő alatt ismert valószínűséggel megtörténő események bekövetkezésének számát (például: egy telefonközpontba adott időszakban és időtartamban beérkezett telefonhívások száma, vagy egy radioaktív anyag adott idő alatt elbomló atomjainak száma).

A Poisson eloszlás a kis valószínűségű, vagy ritka események eloszlása. Poisson eloszlású a sajtóhibák száma egy újság lapjain, egy öntvényben előforduló buborékok száma, egy telefonközpontba beérkező hívások száma, adott idő alatt elbomlott radioaktív atomok száma stb.

Mennyiségek

1. Csatorna kapacitás: adott csatornához a kapacitás- a bináris szimmetrikus csatornával való ekvivalencia alapján:

   Egységnyi idő alatt „átlag” ugyanannyi különböző jelsorozat vihető át, mint egy *C* bit/sec sebességű bináris csatornán. Ekkor a csatorna kapacitás *C*.

2. A forrás szimbólumok sorozatát választja, a választás bizonytalanságát jellemezzük az egy szimbólumra jutó átlagos mennyiséggel, ez lesz a *Shannon-entrópia*.

3. *N* hosszú sorozatok lehetnek, *n* lehetséges szimbólum, $p_1, p_2 \ldots p_n^{N,}$ eloszlás

Mennyiségek

$$1 = \sum_{i=1}^{n^N} p_i$$

p_i-k a valószínűségek

Shannon-entrópia: ξ valószínűségi változó $P(\xi=x_i)=p_i$, az események eloszlása, a ξ valószínűségi változó által felvehető értékek $\{x_1, .., x_i, x_{n^N}\}$, $i=1, \ldots, n^N$

$$\boxed{H(p_1, p_2, .., p_{n^N}) = \mathbb{E}(I(\xi)) = \sum_{i=1}^{n^N} p_i I(\xi = x_i) = -\sum_{i=1}^{n^N} p_i \log_2 p_i}$$

N hosszú forrás egy szimbólumra jutó átlagos bizonytalansága

$$H(p_1, p_2, .., p_n) = \frac{1}{N} \sum_{i=1}^{n^N} p_i \log_2 p_i^{-1}$$

H bit/szimbólum, entrópia; = *1/N H*(az N hosszú üzenetek eloszlása)

Shannon csatorna kódolás alaptétele

C bit/sec csatorna kapacitás, a *H bit/szimbólum* a Shannon-entrópia, *V szimbólum /sec* forrás sebessége

Megválasztható-e a forrás sebessége; hány szimbólumot választ másodpercenként?

Mennyiségi feladat megoldhatósága? *Veszteségmentesen* átvihető-e minden üzenet? Vagyis az átvitt üzenet egyértelműen rekonstruálható-e.

Shannon csatorna kódolás alaptétele (Shannon)

Nem lehet veszteségmentesen működtetni a kommunikációs rendszert ha:

V> C/H ;

Tetszőleges ε> 0-hoz ∃ kódolás: V < C/H- ε sebességgel veszteségmentesen működtethető a csatorna, rendszer

(Ennek formális bizonyítása a legelemibb esetre később).

A csatorna kapacitás

Diszkrét, zajmentes csatorna

$S_1, S_2, \ldots S_m$ a jelek

$t_1, t_2, \ldots t_m$ – átviteli idő jelenként.

Szimmetrikus csatorna $t_1 = t_2 = \ldots = t_m = t$

Szimmetrikus bináris csatorna $m=2,$ a jelek a bitek,{ megint ! *binary digit* (de ez nem ugyanaz mint az információ mértékegysége)}.

Kapacitás a másodpercenként átvihető bitek száma.

Nyilván $C = t^{-1}$, ha t^{-1} egész szám. Ha nem, és t és 1 nem összemérhető, akkor T idő alatt N(T) jelsorozatot lehet, ahol

$$\boxed{2^{T\left(t^{-1}-\varepsilon\right)} < N(T) < 2^{T\left(t^{-1}+\varepsilon\right)}}$$

teljesül tetszőleges ε >0, ha *T* elég nagy

A csatorna kapacitás

A logaritmust képezve (2-es alapú logaritmust használunk a továbbiakban) és T-vel osztva:

$$\boxed{\left(t^{-1}-\varepsilon\right)<\frac{1}{T}log_2 N(T)<\left(t^{-1}+\varepsilon\right)}$$

azaz a határérték $C=t^{-1}$

A csatorna kapacitás

Intuitív:

Két csatorna ekvivalens, ha T hosszú sorozataik száma minden T-re (határértékben) megegyezik. A T hosszú kód-szavak kölcsönösen megfeleltethetők egymásnak

Pontos definíció:

Legyen $N(T)$ egy csatorna T hosszú, (T-nél nem hosszabb, T maximális hosszú) jelsorozatainak ***száma***

A csatorna kapacitás:

$$C = \lim_{T \to \infty} \frac{\log_2 N(T)}{T}$$

Következmény:

Minden C kapacitású csatorna ekvivalens a C kapacitású bináris csatornával. Ezért C dimenziója *bit/sec*.

A csatorna kapacitás

Példa:

1. Szimmetrikus csatorna, *m* db. szimbólum, *t* egy szimbólum átviteli ideje:

$$N(T) \cong m^{T/t}$$

$$C = \lim_{T \to \infty} \frac{\log_2 m^{T/t}}{T} = t^{-1} \log_2 m$$

A csatorna kapacitás

2. Egyszerű nem szimmetrikus csatorna, $t_1, t_2, \dots t_m$ – átviteli idő jelenként; ($S_1, S_2, ..S_m$, szimbólumkészlet; nincs megszorítás a jelsorozatra), $N(T)$-el jelöljük a T *átviteli* idő tartamú sorozatok számát, akkor

$N(T)= N(T- t_1) + N(T- t_2) + \dots + N(T- t_m)$ (*differencia egyenlet*), a sorozatok számának teljes összege (amely sorozatok $S_1, S_2, ..S_m$ szimbólummal/ban végződnek)

A véges differencia egyenletek elmélete alapján $N(T)$ (aszimptotikusan nagy T-kre) jellemezhető

karakterisztikus polinommal: (Shannon-HirkozlesMatElm1948-magy.pdf, 5.o.)

$$\boxed{1 - x^{-t_1} - x^{-t_2} - \dots - x^{-t_m} = 0}$$

A csatorna kapacitás

x_0 : a legnagyobb pozitív megoldás (kis gyökök elhanyagolható adalékokat adnak. $N(T)=k\, x_0^T$

$$x_0^T = \underbrace{(x_0^{-t_1} + x_0^{-t_2} + \ldots + x_0^{-t_m})}_{1} \cdot x_0^T =$$
$$= x_0^{T-t_1} + x_0^{T-t_2} + \ldots + x_0^{T-t_m}$$

x_0 hatványai úgy viselkednek mint $N(T)$.

$$C = \lim_{T \to \infty} \frac{\log_2 N(T)}{T} = \log_2 x_0 = C$$

T lineáris növelés exponenciális növekedést okoz N(T)-ben.

A forrás bizonytalanságának mérőszáma a Shannon-entrópia

A hírközlő rendszer minden lehetséges üzenet választásra (sokszor ismétlődő választásra) működjön. Nem az a fontos az átvitelben, hogy mit választunk, hanem az, hogy milyen *bizonytalan* a választás.

(A szimbólumokat a forrás átnevezheti, ettől független a hírközlés mennyiségi viselkedése)

A forrás N hosszú választásait a n szimbólumból egy n^N elemű **valószínűség-eloszlás** jellemzi.

Az eloszlás meghatározása empirikus alapokon, közelítőleg, történhet. (Később megnézzük az alapeseteket).

A forrás bizonytalanságának mérőszáma a Shannon-entrópia

Feladatunk:

Keressük a diszkrét valószínűség-eloszlások halmazán a

$H(p_1, p_2, \ldots, p_m)$ függvényt, $\Sigma\, p_i=1,\ i=1, \ldots, m,\ p_i \neq 0$;

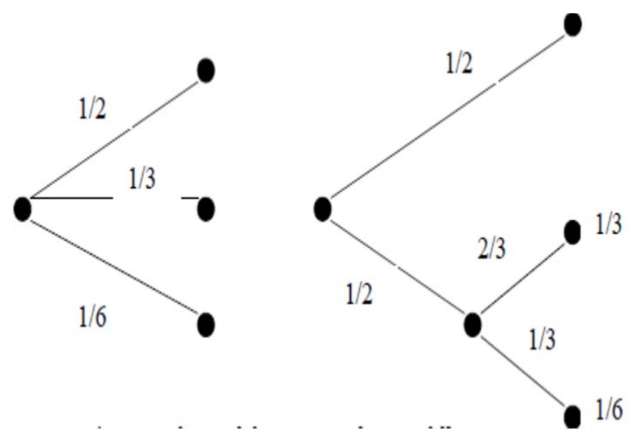
A függvénnyel kapcsolatban támasztott természetes elvárások (Shannon-axiómák):

1. H minden argumentumában legyen folytonos;
2. $A(n)=H(1/n, 1/n, \ldots 1/n) \leq H(1/m, 1/m, \ldots 1/m)=A(m)$, ha $n < m$ (monoton növekedő függvény)
3. Elágazási szabály

A forrás bizonytalanságának mérőszáma a Shannon-entrópia

Elágazási szabály: ha egy választási lehetőség, két rákövetkező választási lehetőségbe ágaztatunk szét. Az eredeti H függvényértékhez hozzá kell adnunk az elágazás valószínűségével súlyozva az elágazás *utáni események feltételes eloszlásának* H függvényét.

$$H(1/2, 1/3, 1/6) = H(1/2, 1/2) + 1/2\, H(2/3, 1/3)$$



A forrás bizonytalanságának mérőszáma a Shannon-entrópia

3. Elágazási szabály (*grouping axiom*) (**ábra**) :

$$H(p_1, \dots p_{n-2}, p_{n-1}, p_n) =$$
$$H(p_1, p_2, \dots, p_{n-1} + p_n) + (p_{n-1} + p_n) H\left(\frac{p_{n-1}}{p_{n-1} + p_n}, \frac{p_n}{p_{n-1} + p_n}\right)$$

Általános elágazási szabály : (indukcióval) legyen,

$\Sigma_{i=1}^{nj} p_{j,i} = q_j$, $i, j=1, \dots, m$, $\Sigma_{j=1}^{m} q_j = 1$, akkor

$$H(p_{1,1}, \dots p_{1,n_1}, p_{2,1}, p_{2,n_2}, \dots, p_{m,1}, \dots, p_{m,n_m}) =$$
$$H(q_1, q_2, \dots, q_m) + \sum_{j=1}^{m} q_j H\left(\frac{p_{j,1}}{q_j}, \dots, \frac{p_{j,n_j}}{q_j}\right)$$

A forrás bizonytalanságának mérőszáma a Shannon-entrópia

Tétel: Az 1-3 axiómákat kielégítő függvény csak a

$K \Sigma p_i \log_2 p_i$, $i=1, \ldots n$ alakú függvények lehetnek, ahol K < 0 konstans.

Bizonyítás:

(i) H(½ , ½)=1, egységnyi entrópia => -*K* (-1)=1 =>

K=-1 – racionális értékű eloszlás esetén.

A forrás bizonytalanságának mérőszáma a Shannon-entrópia

Bizonyítás:

(i) H(½ , ½)=1, egységnyi entrópia => -*K* (-1)=1 =>

K=-1 – racionális értékű eloszlás esetén.

$p_i = q_i/\text{m-re}$, $i=1, \ldots n$, és $\Sigma\, q_i = m$, $\Sigma\, p_i = 1$.

Fejezzük ki $H(p_1, \ldots p_{n-2}, p_{n-1}, p_n)$ –et, $A(m)$ és $A(q_i)$ segítségével.

$$\underbrace{H(p_1, \ldots, p_n)}_{\text{elágazásig}} = \underbrace{H\left(\frac{1}{m}, \ldots, \frac{1}{m}\right)}_{\text{teljes}} - \underbrace{\sum_{i=1}^{n} p_i\, H\left(\frac{1}{q_i}, \frac{1}{q_i}, \ldots, \frac{1}{q_i}\right)}_{\text{elágazás után}} =$$

$$A(m) - \sum_{i=1}^{n} p_i\, A(q_i) = -\sum_{i=1}^{n} p_i\, (A(q_i) - A(m)). \qquad \sum_{i=1}^{n} p_i = 1$$

A forrás bizonytalanságának mérőszáma a Shannon-entrópia

Az *m* elemű egyenletes eloszlásból indulva, csoportok képzése, elágozási szabály felhasználásával

$$A(m) = \underbrace{H(\underbrace{\tfrac{1}{m}, \tfrac{1}{m}, \ldots \tfrac{1}{m}}_{n\ darab})}_{teljes} =$$

$$\underbrace{H\left(\frac{q_1}{m}, \ldots, \frac{q_n}{m}\right)}_{elágazásig} + \underbrace{\sum_{i=1}^{n} \frac{q_i}{m} H\left(\frac{\frac{1}{m}}{\frac{q_i}{m}}, \ldots, \frac{\frac{1}{m}}{\frac{q_i}{m}}\right)}_{elágazás\ utáni} =$$

$$H\left(\frac{q_1}{m}, \ldots, \frac{q_n}{m}\right) + \sum_{i=1}^{n} \frac{q_i}{m} A(q_i)$$

$$H\left(\frac{q_1}{m}, \ldots, \frac{q_n}{m}\right) = H(p_1, \ldots, p_n) = \underbrace{A(m)}_{elágazásig} - \underbrace{\sum_{i=1}^{n} \frac{q_i}{m} A(q_i)}_{elágazás\ utáni} =$$

$$- \sum_{i=1}^{n} \frac{q_i}{m} (A(q_i) - A(m))$$

A forrás bizonytalanságának mérőszáma a Shannon-entrópia

(ii) Lemma: *$A(n) = c \log_2 n$, ahol $c > 0$ tetszőleges*

Biz: Elágazási szabállyal

$$A(nm) = H(\underbrace{\frac{1}{nm}, \ldots, \frac{1}{nm}}_{m}, \ldots, \underbrace{\frac{1}{nm} \ldots, \frac{1}{nm}}_{m}) = H(\underbrace{\frac{1}{n}, \ldots, \frac{1}{n}}_{elágazásig}) + n \cdot \frac{1}{n} H(\underbrace{\frac{1}{m}, \ldots, \frac{1}{m}}_{elágazás \quad után}) = A(n) + A(m)$$

Ebből $A(n^m) = mA(n)$ adódik, mivel A(nm)=A(m)+A(n)

$A(n) = k' \log_2 n$

A forrás bizonytalanságának mérőszáma a Shannon-entrópia

Legyen $s>0$, rögzített egész szám, $\forall$ t-re, és n-re, $\exists$ m,

(1) $s^m < t^n \leq s^{m+1}$;

Az A függvény monoton, $A(s^m) < A(t^n) \leq A(s^{m+1})$ [felhasználva (1)-t].

(2) $mA(s) < nA(t) \leq (m+1)A(s)$ ebből

$$(3) \quad \frac{m}{n} < \frac{A(t)}{A(s)} \leq \frac{m+1}{n} \quad \forall n - re$$

$$(4) \quad \frac{m}{n} < \frac{\log_2 t}{\log_2 s} \leq \frac{m+1}{n} \quad \forall n - re \quad \log_2 \ függvényre$$

A forrás bizonytalanságának mérőszáma a Shannon-entrópia

(3), (4) egyenlőtlenségekből, n-> ∞,

$$\frac{A(t)}{A(s)} = \frac{\log_2 t}{\log_2 s}$$

adódik tehát,

$$C = \frac{A(s)}{\log_2 s} \quad \text{választással } A(t) = C \log_2 t$$

bizonyítja a lemma állítását

A forrás bizonytalanságának mérőszáma a Shannon-entrópia

(iii) (ii) lemma és a racionális elemű eloszlásokra az *(i)* alatt kapott kifejezés együttesen

$$H(p_1,\ \ldots,p_n)= -\sum_{1=1}^{n} p_i\ (A(q_i) - A(m)) = -C\sum_{1=1}^{n} p_i\ \log_2 \frac{q_i}{m} =$$

$$-C\sum_{1=1}^{n} p_i\ \log_2 p_i$$

Ezzel a tételt racionális értékű eloszlásokra igazoltuk.

Tetszőleges eloszlásokra a folytonosságból és *log_2* folytonosságából következik a tétel.

A forrás bizonytalanságának mérőszáma a Shannon-entrópia

H(½ , ½)=1, egységnyi entrópia => -*K* (-1)=1 => K=-1

$$K(\frac{1}{2}\log_2\frac{1}{2}+\frac{1}{2}\log_2\frac{1}{2}) = K(-1) = 1 \Rightarrow K = -1$$

Definició: (p_1, p_2, …, p_m), $\Sigma_{i=1,\ldots,m}\, p_i=1$, *i=1, …,m,* $p_i>0$; eloszlás bizonytalansága a Shannon-entrópia

$$H(p_1,\ \ldots,p_n) = -\sum_{1=1}^{n} p_i\ \log_2 p_i$$

Az entrópia

Az **entrópia** az információ várható értéke:

$$H(p_1, p_2, \ldots, p_m) = \mathbb{E}(I(\xi)) = \sum_{i=1}^{m} p_i I(\xi = x_i) = -\sum_{i=1}^{m} p_i \log_2 p_i$$

Az entrópia tulajdonképpen annak a kijelentésnek a várható információ mennyisége, hogy az *m* db egymást kizáró esemény közül az egyik bekövetkezett.

A $p \log_2 p$ kifejezés $p \to 0$ esetén:

L'Hospital-szabály szerint

$$\lim_{p \to 0} p \log_2 p = \lim_{p \to 0} p \frac{\ln p}{\ln 2} = \frac{1}{\ln 2} \cdot \lim_{p \to 0} \left(\frac{\ln p}{\frac{1}{p}} \right) = \frac{1}{\ln 2} \lim_{p \to 0} \frac{\frac{1}{p}}{\frac{-1}{p^2}} = 0$$

Az entrópia

A forrásunk a következő üzenetet adta le:

„**ccndnddnnncdnndncdncdncdncdnnnncdcdncdcnnncdcccdcddcdccccnnn**”

(21 db „**c**”, 22 db „**n**”, 17 db „**d**”)

$$H(\{c,n,d\}) = 1{,}57$$

Mekkora az üzenet entrópiája?(ha a relatív gyakoriság adja a valószínűség-eloszlást)

$p(\mathbf{c})$=21/60=0,35

$p(\mathbf{n})$=22/60=0,37

$p(\mathbf{d})$=17/60=0,28

$H(\mathbf{c}) = -0{,}35 \log_2 0{,}35 = -0{,}35 \cdot (\ln 0{,}35 / \ln 2) =$
$= -0{,}35 \cdot (-1{,}51) = 0{,}53$

$H(\mathbf{n}) = -0{,}37 \log_2 0{,}37 = -0{,}37 \cdot (\ln 0{,}37 / \ln 2) =$
$= -0{,}37 \cdot (-1{,}43) = 0{,}53$

$H(\mathbf{d}) = -0{,}28 \log_2 0{,}28 = -0{,}28 \cdot (\ln 0{,}28 / \ln 2) =$
$= -0{,}28 \cdot (-1{,}84) = 0{,}51$

Az entrópia

A forrásunk a ♈, ♉, ♊, ♋, ♌ szimbólumokat bocsátja ki egyforma, $p_{♈} = p_{♉} = p_{♊} = p_{♋} = p_{♌} =$ 0,2 valószínűséggel.

Mennyi a forrás, mint halmaz entrópiája?

$$H(♈, ♉, ♊, ♋, ♌) = (-0{,}2 \log_2 0{,}2) \cdot 5 = \underline{2{,}32}$$

Az entrópia

A forrásunk a ♈, ♉, ♊, ♋, ♌ szimbólumokat bocsátja ki

$p_{♈}$=0,12, $p_{♉}$=0,37, $p_{♊}$=0,06, $p_{♋}$=0,21, $p_{♌}$=0,24 valószínűséggel.

Mennyi a forrás, mint halmaz entrópiája?

$$H(♈, ♉, ♊, ♋, ♌) = -0{,}12 \log_2 0{,}12 -$$
$$- 0{,}37 \log_2 0{,}37 - 0{,}06 \log_2 0{,}06 -$$
$$- 0{,}21 \log_2 0{,}21 - 0{,}24 \log_2 0{,}24 =$$
$$= 0{,}37+0{,}53+0{,}24+0{,}47+0{,}49 =$$
$$= \underline{2{,}1}$$

Az entrópia tulajdonságai

1. Nem negatív: $H(p_1, p_2, \ldots, p_m) \geq 0$
2. Az események valószínűségeinek folytonos függvénye.
3. $H(p_1, p_2, \ldots, p_m , 0) = H(p_1, p_2, \ldots, p_m)$
4. Ha $p_i = 1$, a többi $p_k = 0$, ($k=1, \ldots, i-1, i+1 ,\ldots, m$), akkor $H(p_1, p_2, \ldots, p_m) =0$.
5. $H(p_1, p_2, \ldots, p_m) \leq H(1/m, 1/m, \ldots 1/m)$

Egyenlőség csak egyenletes eloszlásra, a legnagyobb bizonytalanságú eloszlásra.

6. $H(p_1, \ldots, p_{k-1},p_\ell ,p_{k+1},\ldots,p_{\ell-1},p_k ,p_{\ell+1},\ldots,p_m) = H(p_1, p_2, \ldots, p_m)$, $\forall\ k, \ell$; azaz az entrópia szimmetrikus változóinak cseréjére.

Az együttes entrópia

Legyen $\xi=\{x_1, \ldots, x_{m_1}\}$ az elsőnek leadott jelek halmaza, $\eta=\{y_1, \ldots, y_{m_2}\}$ pedig a másodiknak leadott jelek halmaza. Vizsgáljuk azt az összetett eseményt, hogy egy ξ-beli és egy η-beli esemény is bekövetkezik.

x_i és y_j együttes bekövetkezési valószínűsége

$$p_{i,j} = p(x_i\ y_j),\ P(\xi=x_i,\ \eta=y_j)$$

a két esemény **együttes bekövetkezésekor nyert információ**

$$I(x_i\ y_j)=-\log_2 p(x_i\ y_j)=-\log_2 p_{i,j}\ .$$

Mindig igaz a együttes információra, hogy $I(x_i\ y_j)\geq I(x_i)$, és $I(x_i\ y_j)\geq I(y_j)$.

A együttes entrópia

a két esemény **együttes bekövetkezésekor nyert információ**

$$I(x_i\ \ y_j)=-\log_2 p(x_i\ \ y_j)=-\log_2 p_{i,j}\ .$$

η és ξ valószínűségi változók **együttes entrópiája**:

$$\boxed{H\left(\xi=x_i,\eta=y_j\right)=-\sum_{i=1}^{m_1}\sum_{j=1}^{m_2} p_{i,j}\log_2 p_{i,j}.}$$

A feltételes entrópia

Legyen $\xi=\{x_1, \ldots, x_{m_1}\}$ a lehetséges leadott jelek halmaza, $\eta=\{y_1, \ldots, y_{m_2}\}$ pedig a vehető jelek halmaza. Minden ξ-beli esemény bekövetkezése maga után von egy η-beli eseményt.

x_i-nek y_j-re vonatkoztatott feltételes valószínűsége $p(x_i | y_j)$. Az ξ halmaz η halmazra vonatkoztatott **feltételes entrópiája**:

$$H(\xi|\eta) = -\sum_{j=1}^{m_2} p(\eta = z_j) \sum_{i=1}^{m_1} p\left(\xi = x_i \middle| \eta = y_j\right) \log_2 p\left(\xi = x_i \middle| \eta = y_j\right) =$$

$$= -\sum_{j=1}^{m_2} \sum_{i=1}^{m_1} p\left(\xi = x_i \middle| \eta = y_j\right) \log_2 p\left(\xi = x_i \middle| \eta = y_j\right).$$

A feltételes entrópia

Legyen $\xi=\{x_1, \ldots, x_{m_1}\}$ a lehetséges leadott jelek halmaza, $\eta=\{y_1, \ldots, y_{m_2}\}$ pedig a vehető jelek halmaza. Minden ξ -beli esemény bekövetkezése maga után von egy η -beli eseményt.

x_i-nek y_j-re vonatkoztatott feltételes valószínűsége $p(x_i | y_j)$.

Mivel $p(x_i y_j)=p(y_j) \cdot p(x_i|y_j)$ minden i-re és j-re,

$$H(\xi\, \eta)= H(\eta) \cdot H(\xi \mid \eta)= H(\xi) \cdot H(\eta \mid \xi).$$

Így $H(\xi) \geq H(\xi\, \eta) \geq 0$

Appendix

Matematikai kitérő –

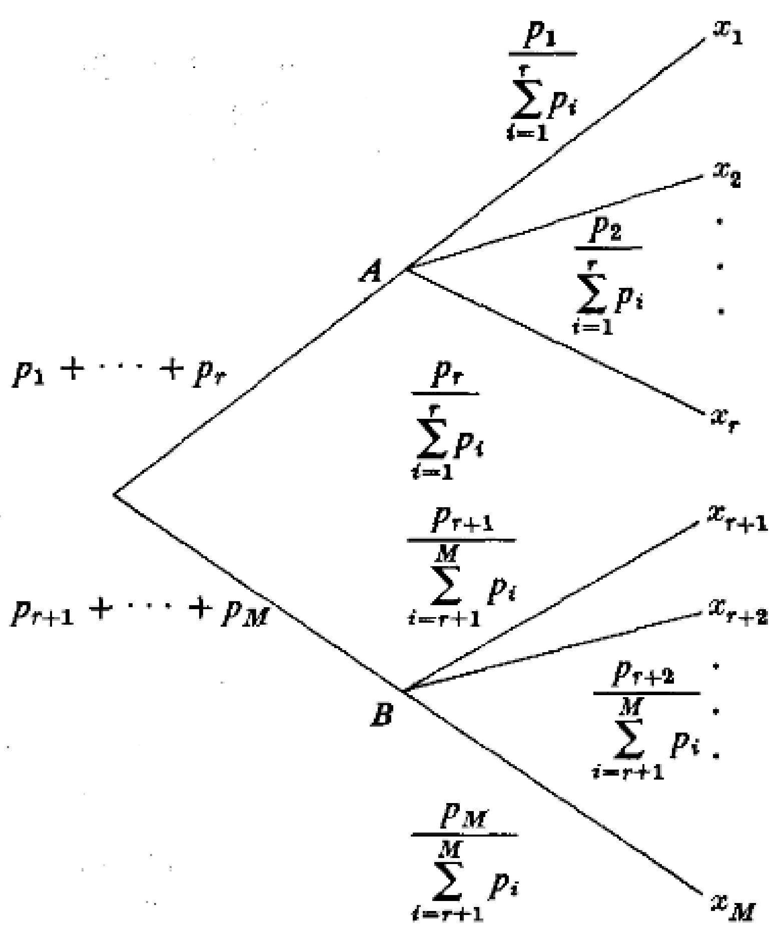
A valószínűségszámítás alapfogalmairól

Egy hírközlési csatorna bemenetén és kimenetén megjelenő jelek nem függetlenek egymástól. Ha *B* jelet vettünk akkor annak a valószínűsége, hogy *A* jel volt a csatorna bemenetén:

A-nak *B* feltétel melletti **feltételes valószínűsége**

$$p(A|B)=\frac{p(A\cdot B)}{p(B)}=\frac{p(A\cap B)}{p(B)}$$

Összetett kisérlet



Elágazás szabály - magyarázat

X valsz. vált. , változó értékeket A és B csoportra bontjuk,

$A - x_1, x_2, \ldots, x_r$; $P(A)=\sum_{i=1}^{r} p_i$

$B - x_{r+1}, x_{r+2}, \ldots, x_M$; $P(B)=\sum_{i=r+1}^{M} p_i$

$P(X=x_i | x_i \in A)=p_i/(\sum_{i=1}^{r} p_i)$;

$P(X=x_i | x_i \in B)=p_i/(\sum_{i=r+1}^{M} p_i)$;

Az összetett kísérlet (ábra) az X valsz. vált. eredeti kísérletével ekvivalens.

Elágazás szabály - magyarázat

Y valsz. vált. , az összetett kísérlet (és annak kimenetei),

$P(Y=x_1)= P$ {*A-t* választottuk, x_1 –t választottuk ki }= P {*A-t* választottuk}, P { x_1 –t választottuk ki | *A-t* választottuk }$=(\sum_{i=1}^{r} p_i)\, p_1/(\sum_{i=1}^{r} p_i)=p_1$;

$\Rightarrow P(Y=x_i) =p_i$ *Y* és *X* valsz. Eloszlása u.a.

Az összetett kísérlet (ábra) az *X* valsz. vált. eredeti kísérletével ekvivalens.

Elágazás szabály - magyarázat

Az összetett kísérlet előtt, az átlagos bizonytalanság: $H(p_1, \dots, p_M)$

Ha feltárjuk, hogy *A* és *B* csoport közül melyiket választottuk, akkor megszüntetünk átlagosan

$$H(p_1 + \dots + p_r, p_{r+1} + \dots + p_M)$$

bizonytalanságot

A-t választottuk ($\sum_{i=1}^{r} p_i$) valószínűséggel, a maradvány bizonytalanság

$$H\left(\frac{p_1}{\sum_{i=1}^{r} p_i}, \frac{p_2}{\sum_{i=1}^{r} p_i}, \dots, \frac{p_r}{\sum_{i=1}^{r} p_i}\right);$$

Elágazás szabály - magyarázat

B-t választottuk ($\sum_{i=r+1}^{M} p_i$) valószínűséggel, a maradvány bizonytalanság

$$H\left(\frac{p_{r+1}}{\sum_{i=r+1}^{M} p_i}, \frac{p_{r+2}}{\sum_{i=r+1}^{M} p_i}, \ldots, \frac{p_M}{\sum_{i=r+1}^{M} p_i}\right);$$

Az átlagos maradvány bizonytalanság valamelyik csoport kiválasztása után

$$(p_1 + \ldots + p_r) H\left(\frac{p_1}{\sum_{i=1}^{r} p_i}, \frac{p_2}{\sum_{i=1}^{r} p_i}, \ldots, \frac{p_r}{\sum_{i=1}^{r} p_i}\right) +$$

$$(p_{r+1} + \ldots + p_M) H\left(\frac{p_{r+1}}{\sum_{i=r+1}^{M} p_i}, \frac{p_{r+2}}{\sum_{i=r+1}^{M} p_i}, \ldots, \frac{p_M}{\sum_{i=r+1}^{M} p_i}\right);$$

Elágazás szabály - magyarázat

A harmadik követelmény H-val (az átlagos bizonytalansággal szemben)

$$H(p_1,\ldots,p_M)=H\left(p_1+\ldots+p_r,p_{r+1}+\ldots+p_M\right)+$$

$$\left(p_1+\ldots+p_r\right)H\left(\frac{p_1}{\sum_{i=1}^{r}p_i},\frac{p_2}{\sum_{i=1}^{r}p_i},\ldots,\frac{p_r}{\sum_{i=1}^{r}p_i}\right)+$$

$$\left(p_{r+1}+\ldots+p_M\right)H\left(\frac{p_{r+1}}{\sum_{i=r+1}^{M}p_i},\frac{p_{r+2}}{\sum_{i=r+1}^{M}p_i},\ldots,\frac{p_M}{\sum_{i=r+1}^{M}p_i}\right);$$

Elágazás szabály - magyarázat

Numerikus példa

$$H(\underbrace{\frac{1}{2},\frac{1}{4}}_{A},\underbrace{\frac{1}{8},\frac{1}{8}}_{B})=H\left(\frac{3}{4},\frac{1}{4}\right)+\frac{3}{4}H\left(\frac{2}{3},\frac{1}{3}\right)+\frac{1}{4}H\left(\frac{1}{2},\frac{1}{2}\right)$$