Entrópia

Ha az eseménytér eseményei nem egyformán valószínűek, akkor a hírkészlet jól jellemezhető a hírek átlagos információ tartalmával.

A *hírkészlet átlagos információtartalmát a hírkészlet entrópiájának* nevezik.

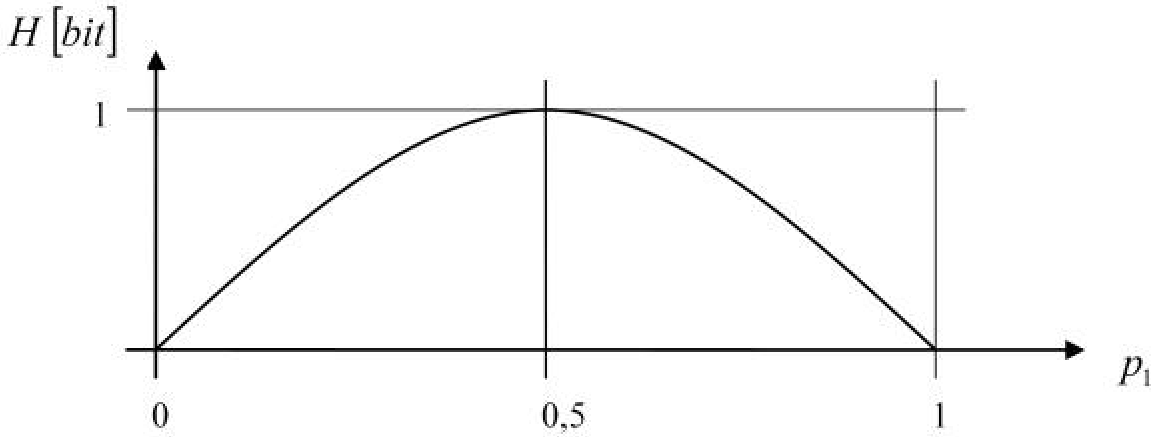
$$ H_E = \sum\limits_{i=1}^n p_i \cdot I_{E_i} = \sum\limits_{i=1}^n p_i \cdot log_2 \frac{1}{p_i} = - \sum\limits_{i=1}^n p_i \cdot log_2 \cdot p_i $$.

Példa

Adott egy eseményrendszer, amely két eseményből áll: $ E = \{E_1, E_2 \} $, továbbá $ p = \{ p_1, p_2\} $ ezért $ p_2 = 1- p_1 $, ekkor az átlagos információtartalom:

$$ H_E = - [p_1 \cdot log p_1 + (1 - p_1) \cdot log (1-p_1)] $$.

Ez a függvény a következőképpen ábrázolható:



Azt láthatjuk, hogy az entrópia akkor a legnagyobb, ha a két esemény egyforma valószínűségű.

Általánosságban elmondható, hogy akkor alacsony az entrópia, amikor kicsi valószínűségű eseményeket is tartalmaz az eseményrendszerünk.

Az entrópia itt arányos a véletlenszerűséggel, ha egy eseményrendszerben magas az entrópia, akkor az egyes elemek előfordulási valószínűségeik közel vannak egymáshoz. Maximális akkor lenne az entrópia, ha az eseményrendszer összes eleme egyformán valószínű.

Példa

Egy négy eseményből álló rendszer előfordulási valószínűségei a következők:

$$ E = \{ E_1, E_2, E_3, E_4\}, $$

az egyes események bekövetkezési valószínűségei a következők:

$$ p = \{ 0.5, 0.25, 0.2, 0.05\} $$.

Mekkora az egyes rendszerállapotokról szóló hírek egyedi információtartalma? Alkalmazzuk az információs összefüggését.

$$ I_{E_1} = - log_2 0.5 = 1 \, [bit] \\ I_{E_2} = - log_2 0.25 = 2 \,[bit] \\ I_{E_3} = - log_2 0.2 = 2.32\, [bit] \\ I_{E_4} = - log_2 0.05 = 4.32\, [bit] \\ $$