Да предположим, че имаме набор от възможни събития, чиито вероятности за настъпване са . Тези вероятности са известни, но това е всичко, което знаем за това кое събитие ще настъпи. Можем ли да намерим мярка за това колко голям е „изборът“ при подбора на събитието или колко несигурни сме относно резултата? Ако има такава мярка, да речем , разумно е да изискваме от нея следните свойства:

1. *H* трябва да е непрекъсната в .

2. Ако всички са равни, следователно , тогава *H* трябва да бъде монотонно нарастваща функция на *n*. При еднакво вероятни събития има повече възможности за избор или несигурност, когато има повече вероятни събития.

3. Ако един избор може да бъде разделен на два последователни избора, то първоначалното *H* трябва да бъде претеглената сума от единичните стойности на *H*.

**Теорема**: Единственото *H*, което изпълнява изброените по-торе три условия, има вида:

където *K* е полижителна константа.

Нека . От условие (3) можем да разделим избора от еднакво вероятни възможности на поредица от *m* избора измежду *s* еднакво вероятни възможности и ще получим . Аналогично . Избираме достатъчно голямо *n* и намираме *m*, което да удовлетворява следното неравенство: . След като логаритмуваме и разделим на , получаваме че:

където е достатъчно малко. О монотонното свойство на получаваме, че:

Следователно, разделяйки на получаваме:

където трябва *K* да бъде положително, за да е изпълнено условие (2).

Нека допуснем, че имаме възможност да избираме измежду *n* възможности с равни вероятности , където са цели числа. Можем да сведем от избор измежду възможности до избор измежду *n* възможности с вероятности и след това, ако е избрана *i*-тата, да направим избор измежду с равни вероятности. Използвайки отново условие (3), приравняваме общия избор измежду , изчислен по два метода

Следователно

Ако са неизмерими, те могат да бъдат приближени с рационални числа, като същото отношение трябва да важи съгласно нашето предположение за непрекъснатост. Следователно отношението важи по принцип. Изборът на коефициента *K* е въпрос на удобство и се свежда до избора на единица мярка.

Величините от вида заемат централно място в теорията на информацията като мярка за информация, избор и несигурност. Изразът *H* се приема като ентропия, така както е дефинирана в някои формулировки на статистическата механика, където е вероятността дадена система да се намира в точка *i* на фазовото си пространство. *H* е същото като *H* в известната теорема на Болцман. Ще наричаме ентропия на множеството от вероятности . Ако *x* е случайна величина, ще записваме *H(x)* като нейна ентропия; по този начин *x* не е аргумент на функция, а обозначение за число, за да се разграничи от *H(y)*, която е ентропията на случайната величина *y*.