Boris Kudryashov

ITMO University

October 18, 2016

1 Prove that the size of the set of all subsets of n-element discrete set is 2^n.

1. Prove that the size of the set of all subsets of n-element discrete set is 2^n.
2. Derive the formula of total probability.

1. Prove that the size of the set of all subsets of n-element discrete set is 2^n.
2. Derive the formula of total probability.
3. Derive the Bayes' formula (Bayes' theorem).

1 Prove that the size of the set of all subsets of n-element discrete set is 2^n.

2 Derive the formula of total probability.

3 Derive the Bayes' formula (Bayes' theorem).

4 Prove the equality $|XY| = |X||Y|$.

5 How many different sequences of length n is it possible to construct from elements of discrete set $X = \{0, ..., q-1\}$?

6 Prove the following properties of numerical characteristics of random variables:

- $M[x + y] = M[x] + M[y]$.
- $M[cx] = cM[x]$.
- $M[xy] = M[x]M[y]$.
- $D[x + y] = D[x] + D[y]$.
- $D[cx] = c^2 D[x]$.
- $D[c + x] = D[x]$.
- If x and y are independent, then $K(x, y) = 0$. That is, independence of random variables implies them to be uncorrelated.

7 For stationary discrete source $X = \{a, b, c\}$ for probability distributions
- $p(a) = p(b) = p(c) = 1/3$;
- $p(a) = p(b) = 1/4$, $p(c) = 1/2$;

calculate the total information and entropy for each of letters. What is the amount of information, contained in the sequence *abaac* ?

8 Prove, thant linear combination of convex functions with positive coeficcients is also a convex function.

8 Prove, thant linear combination of convex functions with positive coeficcients is also a convex function.

9 Prove, that

$$H\left(\frac{\vec{p}_1 + \vec{p}_2 + \vec{p}_3}{3}\right) \geq \frac{H(\vec{p}_1)}{3} + \frac{H(\vec{p}_2)}{3} + \frac{H(\vec{p}_3)}{3}$$

10 Prove the Theorem (1):

Theorem

Let $f(\vec{x})$ be the convex $\cap$ function of vector argument $\vec{x}$, defined on a convex area R and let define constants $\alpha_1, ..., \alpha_M \in [0, 1]$ such that $\sum\limits_{m=1}^{M} \alpha_m = 1$. Then for all $\vec{x}_1, ..., \vec{x}_M \in R$ holds the inequality:

$$f\left(\sum_{m=1}^{M} \alpha_m \vec{x}_m\right) \geq \sum_{m=1}^{M} \alpha_m f(\vec{x}_m). \qquad (1)$$

11 The alphabet of the source is $X = \{a, b, c\}$.
Assume that $p(a) = \alpha$ is defined. Build a
relationship between entropy $H(X)$ and
$p(b) = \beta$.

11 The alphabet of the source is $X = \{a, b, c\}$.
   Assume that $p(a) = \alpha$ is defined. Build a
   relationship between entropy $H(X)$ and
   $p(b) = \beta$.

12 What should relation between parameters α and
   β be, that a Markov chain with the following
   transition matrix:

$$\Pi = \left[\begin{array}{cc} \alpha & 1 - \alpha \\ \beta & 1 - \beta \end{array} \right]$$

   breeds the sequence of independent messages?

13 Without doing calculations, demonstrate stationary probability distribution for the Markov chain with the following transition probability matrix:

$$\Pi = \left[\begin{array}{cc} 1/2 & 1/2 \\ 0 & 1 \end{array} \right].$$

14 *Non-ergodic Markov chain* Find the stationary probability distribution for the Markov chain with the following transition probability matrix:

$$\begin{bmatrix} 1/2 & 0 & 1/2 & 0 \\ 0 & 2/3 & 0 & 1/3 \\ 3/4 & 0 & 1/4 & 0 \\ 0 & 1/2 & 0 & 1/2 \end{bmatrix}.$$

15 Binary source is defined by Discrete-time Markov chain with the following transition probability matrix:

$$\Pi = \left[\begin{array}{cc} 1/2 & 1/2 \\ 1/8 & 7/8 \end{array} \right].$$

Find total information of the sequence 0110, find $H(X)$. Build dependencies $H(X|X^n)$ and $H_n(X)$ from n, find $H(X|X^\infty)$. What impact has "letters dependency" on the speed of creating information by the source?

16 *Presentation of implementation of binary source as a sequence of series lengths.* If we want to store in the memory or in the disk storage binary sequences, containing a small number of ones, we describe them as a sequence of lengths of sequences of zeros. Denote 0^i a sequence of i zeros. As example, the sequence 0011000010100 well be equivalent to (2,0,4,1,2) in this notation. Thus, binary sequence $x_1, ..., x_n, ..., x_i \in X = \{0, 1\}$ will be transformed into the sequence of non-negative integer numbers – $y_1, y_2, ..., y_i \in Y = \{0, 1, ...\}$, which represent length of series of zeros. Let binary source be constant, so that $p(\ "1"\ ) = p$, find the probability distribution of the set of values of the lengths of series Y, entropy $H(Y)$. *Hint:* While solving the problem, you will get the geometric probability distribution, and while entropy calculations, you will need the mathematical expectation formula. See the next problem.

17 *Geometric probability distribution* on set
$I = \{0, 1, 2, ...\}$ is distribution with the following
probabilities: $p(i) = (1 - \alpha)\alpha^i$ (why?). Check
the satisfaction of probability normalization.
Find mathematical expectation, dispersion and
entropy $H(I)$.

- *Hint:* While checking the satisfaction of probability normalization, you should calculate the sum like:

$$\sum_{i=0}^{\infty} a^i. \tag{2}$$

Remember the geometric progression sum formulae. To calculate the mathematical expectation, you should deal with the sum like:

$$\sum_{i=0}^{\infty} i a^i. \tag{3}$$

Note, that if we differentiate (2) by a, we get almost (3). Thus, we have reduced the problem to the differentiation of well known geometric progression sum formulae. The second derivative of this expression should be used for dispersion calculation. One more hint for calculating sum (3) is, that the factor i under the sum sign can be represented as the sum i ones, and then the order of the summation is changed. *Ответы:*

$$\mathbf{M}[i] = \frac{\alpha}{1-\alpha}; \quad \mathbf{D}[i] = \frac{\alpha}{(1-\alpha)^2}; \quad H(I) = \frac{\eta(\alpha)}{1-\alpha}.$$

18 *Relative entropy of Kullback–Leibler of distribution $\{p(x), x \in X\}$ according to distribution $\{q(x), x \in X\}$ is the following function:*

$$L(p||q) = \sum_{x \in X} p(x) \log \frac{p(x)}{q(x)}. \qquad (4)$$

This function should be considered as the the measure of difference of two distributions. Indeed, the relative entropy is equal to zero, if this two distributions are equal, and is strictly positive if they differ. Prove the latter statement.

Hint. Rewrite the relative entropy like:

$$L(p||q) = -\sum_{x \in X} p(x) \log \frac{q(x)}{p(x)}.$$

and use the following inequality $\log z \leq (z - 1) \log e$.

19 Uniform distribution is *extreme* meaning that it has the largest entropy among all probability distributions on the set of finite fixed size. Prove the following statement: *Geometric distribution is extreme in (maximizing the entropy) for discrete random values, which take Countable set of values and have specified mathematical expectation.*

- *Solution.* Let $p(i) = (1 - \alpha)\alpha^i, i = 0, 1, \ldots$ – geometric distribution, $q(i)$ – arbitrary distribution on the same set of numbers with the same mathematical expectation.
  $\mathbf{M}[i] = \alpha/(1 - \alpha)$, $H(p)$ и $H(q)$ – corresponding entropy's, such that $H(p) = \eta(\alpha)/(1 - \alpha)$ (see the 17-th pronlem). Then,

$$
\begin{aligned}
H(q) &= -\sum_{i=0}^{\infty} q(i) \log q(i) = -\sum_{i=0}^{\infty} q(i) \log \left(\frac{q(i)}{p(i)} p(i) \right) = \\
&= -L(p\|q) - \sum_{i=0}^{\infty} q(i) \log p(i) \leq \\
&\leq -\sum_{i=0}^{\infty} q(i) \left(\log(1 - \alpha) + i \log \alpha \right) = \\
&= -\log(1 - \alpha) - \mathbf{M}[i] \log \alpha = \\
&= -\log(1 - \alpha) - \frac{\alpha \log \alpha}{1 - \alpha} = H(p).
\end{aligned}
$$

20 Consider a constant source of ternary messages. The alphabet of the source is $X = \{a, b, c\}$. Using the following probability distributions,

❶ $p(a) = p(b) = p(c) = 1/3$;
❷ $p(a) = p(b) = 1/4$, $p(c) = 1/2$;

build the dependency of *uniform code rate* and length of encoded blocks with zero error probability. Compare the code rate with the entropy of the source.

20 Consider a constant source of ternary messages. The alphabet of the source is $X = \{a, b, c\}$. Using the following probability distributions,

  ❶ $p(a) = p(b) = p(c) = 1/3$;
  ❷ $p(a) = p(b) = 1/4$, $p(c) = 1/2$;

build the dependency of *uniform code rate* and length of encoded blocks with zero error probability. Compare the code rate with the entropy of the source.

21 Consider ternary sources from the previous problem. Build the dependency between code rate of uniform code and length of encoded sequences with the error probability equal to 0.1.

22 Hamming weight of sequences $\vec{x}$ over the alphabet $\{0, \ldots, q-1\}$ is the number of non-zero elements in $\vec{x}$. How many binary sequences of length n and w exist?

- *Hint:* You should show, that among all possible *Combinations* of w elements of the set of n different elements, taking into account the order order is equal to:

$$A_n^w = n(n-1) \times \cdots \times (n-w+1).$$

After, you should show, that the number of different permutations of w elements is equal to

$$P_w = w(w-1) \times \cdots \times 1 = w!.$$

At last, number of sequences of length n and weight w is equal to the number of *Combinations* of w elements from n elements ignoring the order of elements, and is equal to:

$$C_n^w = \binom{n}{w} = \frac{A_n^w}{P_w} = \frac{n(n-1) \times \cdots \times (n-m+1)}{w!} =$$

$$= \frac{n!}{w!(n-w)!}.$$

23 Find the probability $p_n(w)$, that a sequence of length n at the output of binary constant source, where the probability of 1 is p, will have Hamming weight equal to w.

Сщььуте: Answer to this problem is the binomial distribution formula.

$$p_n(w) = \binom{n}{w} p^w (1-p)^{n-w}.$$

24 Binary constant source $X = \{0, 1\}$ breeds ones with probability $p = 0,02$. Consider the dependence between error probability and length of encoded blocks, with the code rate 0.5 bit for a letter of source.

24 Binary constant source $X = \{0, 1\}$ breeds ones with probability $p = 0,02$. Consider the dependence between error probability and length of encoded blocks, with the code rate 0.5 bit for a letter of source.

25 Find the probability that sequence with length $n = 10$ at the output of binary constant source, where probability of "1" is $p = 0,1$, will have Hamming weight greater or equal to $w = 5$. (Use your PC). Compare this result to the Chebyshev's estimation of probability of this event.

26 *Polynomial distribution.* How many ternary sequences of length n contain τ_0 zeros, τ_1 ones and $\tau_2 = n - \tau_0 - \tau_1$ twos? Find the probability of such sequence with fixed probability distribution on elements of ternary alphabet. Generalize the result to the alphabets of arbitrary size.

Hint: The answer can be fount in the chapter "Universal Coding".